GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 10, 2005, 18:45:41 ; Search time 182 Seconds
(without alignments)
2841.201 Million cell updates/sec

Title: US-09-516-728A-4
Perfect score: 7030
Sequence: 1 MKPAAREARLPPRSPGLRWA.....TIYENLAPVTTFGKTNGYIA 1337

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7030	100.0	1337	7	ADL16193	Adl16193 Human pro
2	7030	100.0	1337	8	ADP74603	Adp74603 Amino aci
3	6950	98.9	1337	7	ADL16195	Adl16195 Human pro
4	6824	97.1	1337	2	AAR85203	Aar85203 huDEP-1.
5	4508.5	64.1	1238	7	ADL16198	Adl16198 Mouse pro
6	4499.5	64.0	1216	7	ADL16200	Adl16200 Rat prote
7	1811	25.8	341	8	ADP74604	Adp74604 Amino aci
8	1259.5	17.9	1998	7	AAO24268	Aao24268 Murine va
9	1242	17.7	1997	3	AAB19774	Aab19774 Human pro
10	1242	17.7	1997	4	AAM78821	Aam78821 Human pro
11	1242	17.7	1997	8	ADO04580	Ado04580 Human HPT
12	1241	17.7	1997	5	AAE20278	Aae20278 Human lun
13	1241	17.7	1997	7	ABR57179	Abr57179 Human PTP
14	1241	17.7	1997	7	ADJ70329	Adj70329 Human hea
15	1241	17.7	2002	4	AAM79805	Aam79805 Human pro
16	1230	17.5	1450	7	ABR57182	Abr57182 Human PTP
17	1177.5	16.7	1115	8	ADN12200	Adn12200 Human Sap
18	1140	16.2	1128	8	ABM82538	Abm82538 Human dia
19	1130	16.1	1118	6	ABR43688	Abr43688 Human tyr
20	1127	16.0	1093	6	ABR43689	Abr43689 Human tyr
21	1124.5	16.0	1099	8	ABM82535	Abm82535 Human dia
22	1112	15.8	1102	8	ADE28325	Ade28325 Human KPP
23	1107.5	15.8	1083	8	ABM82537	Abm82537 Human dia
24	1105	15.7	1128	8	ABM82533	Abm82533 Human dia
25	1101	15.7	1531	7	ADE08753	Ade08753 Novel pro
26	1088	15.5	1112	8	ABM82534	Abm82534 Human dia
27	1079	15.3	1084	8	ABM82536	Abm82536 Human dia
28	1043.5	14.8	1647	4	ABB58428	Abb58428 Drosophil
29	993.5	14.1	579	3	AAB19773	Aab19773 Mouse vas
30	991.5	14.1	1705	5	ABR52348	Abr52348 Protein r
31	969	13.8	1767	4	ABB62726	Abb62726 Drosophil
32	957.5	13.6	1711	5	ABR52349	Abr52349 Protein r
33	954	13.6	1711	2	AAW70506	Aaw70506 Osteotest
34	944	13.4	1711	2	AAW70507	Aaw70507 Mutant os
35	924.5	13.2	1447	7	ABR57181	Abr57181 Drosophil
36	889.5	12.7	336	8	ADO04582	Ado04582 Human HPT
37	879	12.5	1912	8	ADR40183	Adr40183 Human pro
38	878.5	12.5	312	8	ADO04584	Ado04584 Human HPT
39	878.5	12.5	319	8	ADO04585	Ado04585 Human HPT
40	864	12.3	2301	6	ABP60058	Abp60058 Human pho
41	855	12.2	1188	2	AAW42991	Aaw42991 Amino aci
42	854	12.1	1188	5	ABR52338	Abr52338 Protein r
43	854	12.1	1188	7	ADD89075	Add89075 TAT279. 1
44	854	12.1	1188	8	ADS17078	Ads17078 Human 207
45	849.5	12.1	1217	7	ADD46195	Add46195 Rat Prote

ALIGNMENTS

RESULT 1
ADL16193
ID ADL16193 standard; protein; 1337 AA.
XX
AC ADL16193;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human protein tyrosine phosphatase #13.
XX
KW cytostatic; immunosuppressive; antiallergic;
KW protein tyrosine phosphatase; reversible oxidation; dephosphorylation;
KW inducible signalling pathway; cell proliferation; cancer;
KW guest vs. host disease; autoimmune disease; allergy; metabolic disorder;
KW cell-cycle abnormality; enzyme.
XX
OS Homo sapiens.
XX
PN WO2003068984-A2.
XX
PD 21-AUG-2003.
XX
PF 13-FEB-2003; 2003WO-EP001446.
XX
PR 13-FEB-2002; 2002US-0356810P.
PR 12-FEB-2003; 2003US-00366547.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
PA (CEPT-) CEPTYR INC.
XX
PI Tonks NK, Tzu-Ching M, Cool DE;
XX
DR WPI; 2003-712572/67.
DR N-PSDB; ADL16192.
XX
PT Identifying reversibly oxidized protein tyrosine phosphatase, useful in
PT screening for specific modulators, potential agents for treating e.g.
PT cancer or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 42; 238pp; English.
XX
CC The invention relates to a method for identifying a protein tyrosine
CC phosphatase (PTP) that is reversibly oxidized in a cell by: (i)
CC subjecting a sample, including a cell that contains at least one PTP, to
CC conditions that cause reversible oxidation of PTP; (ii) isolating PTP
CC anaerobically, in presence of a sulfhydryl-reactive agent (II) that
CC irreversibly modifies the thiol group of an invariant Cys in the active
CC site of PTP; and (iii) determining, under reducing conditions, the level

CC of dephosphorylation, caused by PTP, of a labelled substrate (III), where
CC dephosphorylation indicates that an active PTP is present. . No details
CC of tests for these activities are given. The method is used to identify
CC reversibly oxidized PTP, also to identify agents that: (a) reversibly
CC modify such PTP; or (b) alter inducible signalling pathways in which PTP
CC are involved. These agents are potentially useful, in human or veterinary
CC medicine, for treating abnormal cell proliferation or growth (cancer);
CC guest vs. host disease; autoimmune diseases; allergy or other
CC immunosuppressed states; metabolic disorders and cell-cycle
CC abnormalities. This sequence represents one of the PTP enzyme of the
CC invention.
XX
SQ Sequence 1337 AA;

Query Match 100.0%; Score 7030; DB 7; Length 1337;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337
```

RESULT 2
ADP74603
ID ADP74603 standard; protein; 1337 AA.
XX
AC ADP74603;
XX
DT 26-AUG-2004 (first entry)
XX
DE Amino acid sequence of human DEP-1.
XX
KW human; density enhanced phosphatase-1; DEP-1; DEP-1 complex;
KW p120 catenin; adaptor protein; Gab 1; hepatocyte growth factor receptor;
KW HGF receptor; scatter factor receptor; SF receptor; Met; plakoglobin;
KW beta-catenin; signal transduction; cell growth; cell proliferation;
KW cell cycle regulation; contact inhibition; cellular differentiation;
KW cellular morphogenesis; cellular motility; cytoskeleton.
XX
OS Homo sapiens.
XX
PN WO2004048549-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038089.
XX
PR 26-NOV-2002; 2002US-0429746P.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
XX
PI Palka-Hamblin HL, Tonks NK;
XX
DR WPI; 2004-450367/42.
DR N-PSDB; ADP74602.
XX
PT New isolated complex comprises Density Enhanced Phosphatase-1 (DEP-1)

```
PT   polypeptide, useful for manipulating biological signal transduction
PT   pathways, or determining additional molecular components of the pathways.
XX
PS   Claim 2; Page 95-97; 130pp; English.
XX
CC   The present sequence represents a human density enhanced phosphatase-1
CC   (DEP-1) polypeptide. The specification describes DEP-1 complexes,
CC   comprising DEP-1 and DEP-1 substrates. DEP-1 substrate polypeptides that
CC   interact specifically with the substrate-trapping mutant form of DEP-1
CC   are isolated from human breast tumour cell lines. These polypeptides
CC   include the functional component of p120 catenin, the adaptor protein Gab
CC   1, the hepatocyte growth factor (HGF)/scatter factor (SF) receptor Met, a
CC   plakoglobin polypeptide, and a beta-catenin polypeptide. The DEP-1
CC   complexes are useful for manipulating biological signal transduction
CC   pathways, including defining therapeutic targets, or determining
CC   additional molecular components of such pathways. Agents which modulate
CC   the complexes are useful for therapeutic interventions which influence
CC   biological processes, e.g. cell growth or proliferation including cell
CC   cycle regulation and contact inhibition of cell growth, cellular
CC   differentiation including altered cellular morphogenesis or motility or
CC   other cellular activities characterized by alterations in cytoskeletal
CC   organization and/or cellular gene expression, or cell survival including
CC   cellular responses to apoptotic stimuli.
XX
SQ   Sequence 1337 AA;

  Query Match             100.0%;  Score 7030;  DB 8;  Length 1337;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1337;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337


RESULT 3
ADL16195
ID   ADL16195 standard; protein; 1337 AA.
XX
AC   ADL16195;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human protein tyrosine phosphatase #14.
XX
KW   cytostatic; immunosuppressive; antiallergic;
KW   protein tyrosine phosphatase; reversible oxidation; dephosphorylation;
KW   inducible signalling pathway; cell proliferation; cancer;
KW   guest vs. host disease; autoimmune disease; allergy; metabolic disorder;
KW   cell-cycle abnormality; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO2003068984-A2.
XX
PD   21-AUG-2003.
XX
PF   13-FEB-2003; 2003WO-EP001446.
```

XX
PR 13-FEB-2002; 2002US-0356810P.
PR 12-FEB-2003; 2003US-00366547.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
PA (CEPT-) CEPTYR INC.
XX
PI Tonks NK, Tzu-Ching M, Cool DE;
XX
DR WPI; 2003-712572/67.
DR N-PSDB; ADL16194.
XX
PT Identifying reversibly oxidized protein tyrosine phosphatase, useful in
PT screening for specific modulators, potential agents for treating e.g.
PT cancer or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 44; 238pp; English.
XX
CC The invention relates to a method for identifying a protein tyrosine
CC phosphatase (PTP) that is reversibly oxidized in a cell by: (i)
CC subjecting a sample, including a cell that contains at least one PTP, to
CC conditions that cause reversible oxidation of PTP; (ii) isolating PTP
CC anaerobically, in presence of a sulfhydryl-reactive agent (II) that
CC irreversibly modifies the thiol group of an invariant Cys in the active
CC site of PTP; and (iii) determining, under reducing conditions, the level
CC of dephosphorylation, caused by PTP, of a labelled substrate (III), where
CC dephosphorylation indicates that an active PTP is present. . No details
CC of tests for these activities are given. The method is used to identify
CC reversibly oxidized PTP, also to identify agents that: (a) reversibly
CC modify such PTP; or (b) alter inducible signalling pathways in which PTP
CC are involved. These agents are potentially useful, in human or veterinary
CC medicine, for treating abnormal cell proliferation or growth (cancer);
CC guest vs. host disease; autoimmune diseases; allergy or other
CC immunosuppressed states; metabolic disorders and cell-cycle
CC abnormalities. This sequence represents one of the PTP enzyme of the
CC invention.
XX
SQ Sequence 1337 AA;

Query Match 98.9%; Score 6950; DB 7; Length 1337;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1325; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISGLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              |||||||||||||||||    :    |  ||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYLQWEAGTSGLLPACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337
```

RESULT 4
AAR85203
ID AAR85203 standard; protein; 1337 AA.
XX
AC AAR85203;
XX
DT 12-FEB-1996 (first entry)
XX
DE huDEP-1.
XX

```
KW   Density enhanced Type III receptor-like protein tyrosine phosphatase;
KW   huDEP-1.
XX
OS   Homo sapiens.
XX
PN   WO9530008-A1.
XX
PD   09-NOV-1995.
XX
PF   03-MAY-1995;   95WO-US005512.
XX
PR   03-MAY-1994;   94US-00237940.
XX
PA   (COLD-) COLD SPRING HARBOR LAB.
XX
PI   Tonks NK,  Oestman A;
XX
DR   WPI; 1995-393079/50.
DR   N-PSDB; AAT06027.
XX
PT   New density enhanced protein tyrosine phosphatase - used to develop
PT   prods. to modify transcription, translation and activity of tyrosine
PT   phosphatase(s).
XX
PS   Claim 4; Page 34-38; 51pp; English.
XX
CC   A cDNA clone was obtd. (see AAT06027) from a HeLa cell cDNA library that
CC   encoded a novel density-enhanced Type III receptor-like PTP, designated
CC   huDEP-1 (AAR85203). huDEP-1 is useful for the study of PTPs and for the
CC   development of therapeutic or prophylactic cpds. e.g. for prevention of
CC   abnormal or malignant cell growth
XX
SQ   Sequence 1337 AA;

  Query Match             97.1%;  Score 6824;  DB 2;  Length 1337;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches 1302;  Conservative    6;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVAHGCEEGCSLSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLAQKVAWMPAIQREAG 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              |:|  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 QGAPPPLCMMSPFVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVELFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337


RESULT 5
ADL16198
ID   ADL16198 standard; protein; 1238 AA.
XX
AC   ADL16198;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Mouse protein tyrosine phosphatase #4.
XX
KW   cytostatic; immunosuppressive; antiallergic;
KW   protein tyrosine phosphatase; reversible oxidation; dephosphorylation;
KW   inducible signalling pathway; cell proliferation; cancer;
KW   guest vs. host disease; autoimmune disease; allergy; metabolic disorder;
KW   cell-cycle abnormality; enzyme.
```

XX
OS Mus musculus.
XX
PN WO2003068984-A2.
XX
PD 21-AUG-2003.
XX
PF 13-FEB-2003; 2003WO-EP001446.
XX
PR 13-FEB-2002; 2002US-0356810P.
PR 12-FEB-2003; 2003US-00366547.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
PA (CEPT-) CEPTYR INC.
XX
PI Tonks NK, Tzu-Ching M, Cool DE;
XX
DR WPI; 2003-712572/67.
DR N-PSDB; ADL16197.
XX
PT Identifying reversibly oxidized protein tyrosine phosphatase, useful in
PT screening for specific modulators, potential agents for treating e.g.
PT cancer or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 47; 238pp; English.
XX
CC The invention relates to a method for identifying a protein tyrosine
CC phosphatase (PTP) that is reversibly oxidized in a cell by: (i)
CC subjecting a sample, including a cell that contains at least one PTP, to
CC conditions that cause reversible oxidation of PTP; (ii) isolating PTP
CC anaerobically, in presence of a sulfhydryl-reactive agent (II) that
CC irreversibly modifies the thiol group of an invariant Cys in the active
CC site of PTP; and (iii) determining, under reducing conditions, the level
CC of dephosphorylation, caused by PTP, of a labelled substrate (III), where
CC dephosphorylation indicates that an active PTP is present. . No details
CC of tests for these activities are given. The method is used to identify
CC reversibly oxidized PTP, also to identify agents that: (a) reversibly
CC modify such PTP; or (b) alter inducible signalling pathways in which PTP
CC are involved. These agents are potentially useful, in human or veterinary
CC medicine, for treating abnormal cell proliferation or growth (cancer);
CC guest vs. host disease; autoimmune diseases; allergy or other
CC immunosuppressed states; metabolic disorders and cell-cycle
CC abnormalities. This sequence represents one of the PTP enzyme of the
CC invention.
XX
SQ Sequence 1238 AA;

Query Match 64.1%; Score 4508.5; DB 7; Length 1238;
Best Local Similarity 67.2%; Pred. No. 1.3e-292;
Matches 901; Conservative 108; Mismatches 226; Indels 105; Gaps 12;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              |||||||  | ||||||||||||  ||||||  ||:||||  |:||
Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLRQGQVLCAGAAPNPI---------------- 43

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         44 ------------------------------------------------------------ 43

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
               |||:|| :|||:|:|||| ||: ||| :  :::|||:  |  | :|  ||||||  | |
Db         44 --FDIEAV-VSPTSVLLTWKHNDSGASECR--IENKMESNLTFPVKNQTSCNITGLSPGT 98

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              || |||     |||        : ||| ||||| |   || :| |:||| ||||| |:|:
Db         99 SYTFSIISVTTNETLNK----TITTEPWPVSDLHVTSVGVTQARLTWSNANGTASYRMLI 154

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLL-QSNKTKGDPLGTEGGLDASNTERS 299
              | : :|          |||| ||||     | :   :||:|: |    |   ||: |:|
Db        155 EELTTHS--------SVNISGLKPGTN---NTFAFPESNETQADFAVAEEVPDANGTKRI 203

Qy        300 RAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQA 359
                  : : :|  |||  ||| |||    ||:||  |:| |:||||:||||||||||||:

Db        204 PV-TNLSQLHKNSLVS-VDPPSGQDPSLTEILLTDLKPDTQYNATIYSQAANGTEGQPRN 261

Qy        360 IEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419
                |:||: || || |:||||:|:|| || : : | ::  |||||| | | |  |::  |
Db        262 KVFKTNSTQVSDVRAMNISASSMTLTWKSNYDGSRTSIVYKIHVAGGTHSVNQTVNKTEA 321

Qy        420 VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA 479
              :| || ||| ||||| | ||  |||||||||:| |  |||||| |||  ||||| |:|:
Db        322 IILGLSSSTLYNITVHPFLGQTEGTPGFLQVYTSPDQVSDFRVTNVSTRAIGLAWRSNDS 381

Qy        480 ESFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTV 539
              :||:: | |:|    |   | ||| ::  | ||| | |||:|:||:|||| | ||   |
Db        382 KSFEIFIKQDGGEKHRNASTGNQSYMVEDLKPGTSYHFEIIPRGPDGTEGLSSTVNGSTD 441

Qy        540 PSAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGS-NHTSTYDKAITLQGLIP 598
              |||| || || ::|||| |:|:: | || | ||||:||| ||    |::  | ||:  | |
Db        442 PSAVTDIRVVNISTTEMQLEWQNTDDASGYTYHLVLESKSGSIIRTNSSQKWITVGSLTP 501

Qy        599 GTLYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYC 658
              |||||:|| |||| : |  ||  ||||||:||:|:|:| || | : |:| | || :|:|
Db        502 GTLYNVTIFPEVDQIQGISNSITQYTRPSSVSHIEVNTTTTTAAIRWKNEDAASASYAYS 561

Qy        659 LLIEKAGNSSNATQVVT-DIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCT 717
              :|| | |: || |  | |  |   : ||||| ||||:|  |||||  || ||   |||
Db        562 VLILKTGDGSNVTSNFTKDPSI---LIPELIPGVSYTVKILTQVGDGTTSLVPGWNLFCT 618

Qy        718 DPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYR 777
              :|  : || |||||||||||||| || |  |||| | | :|:| | ||:|:|:: || |
Db        619 EPEPVTSFHCEVVPKEPALVLKWACPFGMYTGFELGVRSDSWDNMTRLENCTSDDDTECR 678

Qy        778 TEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKF 837
              ||| |||||||||||| :|:|||||| | :| |||||||||  |||||||||||||||||
Db        679 TEVAYLNFSTSYNISIATLSCGKMALPAQNICTTGITDPPTPDGSPNITSVSHNSVKVKF 738

Qy        838 SGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQS 897
              |||||||||||||||||||||  ||||||||||:|||:|||||||||||| ||||:||
Db        739 SGFEASHGPIKAYAVILTTGEAAQPSADVLKYTYEDFKRGASDTYVTYLIRIEEKGQSQG 798

Qy        898 LSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSF 957
              ||||| |||||||:|||||||||:|||||||||||||||||:: || |||:| ||||||
Db        799 LSEVLNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYVSF 858

Qy        958 SRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKK 1017
              | ||:|| |||||||||||||||| | ||||||||||| ||||||||||||||||||||
Db        859 SPYSEAVFLPQDPGVICGAVFGCIFGALAITAVGGFIFWRKKRTDAKNNEVSFSQIKPKK 918

Qy       1018 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPY 1077
              |||||||||||||||||||||||||||||||:||| ||| ||:|||||||||||||||||
Db        919 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYTAEIAENRGKNRYNNVLPY 978

Qy       1078 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIM 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        979 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIVM 1038

Qy       1138 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQ 1197
              |||||||||||||||||||||||||||||||||||:|||||||| |||:| ||||||||
Db       1039 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEVVLPEWTIRDFVVKNMQNSESHPLRQ 1098

Qy       1198 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIY 1257
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1099 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRLIY 1158

Qy       1258 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAM 1317
              |||||||||||||||||||||||||||||||||||||||||:|:|||||||||||||||
Db       1159 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTTAM 1218

Qy       1318 TIYENLAPVTTFGKTNGYIA 1337
              |||||| ||: |||||||||
Db       1219 TIYENLEPVSMFGKTNGYIA 1238
```

RESULT 6
ADL16200

```
ID   ADL16200 standard; protein; 1216 AA.
XX
AC   ADL16200;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Rat protein tyrosine phosphatase #4.
XX
KW   cytostatic; immunosuppressive; antiallergic;
KW   protein tyrosine phosphatase; reversible oxidation; dephosphorylation;
KW   inducible signalling pathway; cell proliferation; cancer;
KW   guest vs. host disease; autoimmune disease; allergy; metabolic disorder;
KW   cell-cycle abnormality; enzyme.
XX
OS   Rattus norvegicus.
XX
PN   WO2003068984-A2.
XX
PD   21-AUG-2003.
XX
PF   13-FEB-2003; 2003WO-EP001446.
XX
PR   13-FEB-2002; 2002US-0356810P.
PR   12-FEB-2003; 2003US-00366547.
XX
PA   (COLD-) COLD SPRING HARBOR LAB.
PA   (CEPT-) CEPTYR INC.
XX
PI   Tonks NK,  Tzu-Ching M,  Cool DE;
XX
DR   WPI; 2003-712572/67.
DR   N-PSDB; ADL16199.
XX
PT   Identifying reversibly oxidized protein tyrosine phosphatase, useful in
PT   screening for specific modulators, potential agents for treating e.g.
PT   cancer or autoimmune disease.
XX
PS   Disclosure; SEQ ID NO 49; 238pp; English.
XX
CC   The invention relates to a method for identifying a protein tyrosine
CC   phosphatase (PTP) that is reversibly oxidized in a cell by: (i)
CC   subjecting a sample, including a cell that contains at least one PTP, to
CC   conditions that cause reversible oxidation of PTP; (ii) isolating PTP
CC   anaerobically, in presence of a sulfhydryl-reactive agent (II) that
CC   irreversibly modifies the thiol group of an invariant Cys in the active
CC   site of PTP; and (iii) determining, under reducing conditions, the level
CC   of dephosphorylation, caused by PTP, of a labelled substrate (III), where
CC   dephosphorylation indicates that an active PTP is present. . No details
CC   of tests for these activities are given. The method is used to identify
CC   reversibly oxidized PTP, also to identify agents that: (a) reversibly
CC   modify such PTP; or (b) alter inducible signalling pathways in which PTP
CC   are involved. These agents are potentially useful, in human or veterinary
CC   medicine, for treating abnormal cell proliferation or growth (cancer);
CC   guest vs. host disease; autoimmune diseases; allergy or other
CC   immunosuppressed states; metabolic disorders and cell-cycle
CC   abnormalities. This sequence represents one of the PTP enzyme of the
CC   invention.
XX
SQ   Sequence 1216 AA;

  Query Match             64.0%;  Score 4499.5;  DB 7;  Length 1216;
  Best Local Similarity   67.1%;  Pred. No. 5.2e-292;
  Matches  900;  Conservative  106;  Mismatches  205;  Indels  131;  Gaps   10;

Qy          1 MKPAARERARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||| | |||||||||||  |||||| ||::| |
Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLRQGQVVCTG---------------------- 37

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
                                                                  ||
Db         38 ------------------------------------------------------AAPS 41

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||::||: |||:|:|||| ||:| ||||        |    || :|   ||||| |||
Db         42 PVFDVEAVT-SPTSVVLTWKHNDSATSEYKI----NEGNTLRYTVKNQTSFNITGLSPAT 96

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              || |||| |  ||| | |    : ||| |||||:||  || :| |:||| ||||| |: :
Db         97 SYKFSITLGTVNETSGKPTYKNITTEPWPVSDLQVAYIGVTQALLAWSNANGTASYRMQI 156

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
                  ||| :|     |||||||   ::   :  ||:|: |   ||
Db        157 ------VELTTNS--SGGISDLKPGTHKSL--AVQGSNETQHDLWVTE------------ 194

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
                        |  || | :    |||||  |:| |:|  |:||||||||||||
Db        195 --------------GVSDPPSARDPSLTEVLLTELKPDTQYKVTIYSQAADGTEGQPGNK 240

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              |:|| ||| |: ||||| :::|| || ::||| :::||||:||| :|| |  |:| :||
Db        241 VFKTNPIQVSDIRAVNISDSNMTLTWKSNNNESHASFTYKIYVAGGSDSINETVNETQAV 300

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              | || ||| ||||| | ||   | ||||||:| | |||||||| ||  |||||| |:|:|
Db        301 IRGLSSSTLYNITVLPFLGQTAGIPGFLQVYTSPRPVSDFRVTNVSLREIGLAWRSNDSE 360

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||:: |||||:       | : | |:  | ||| | ||| |:|||||| |:||  ||
Db        361 SFEIFITQEGSEKRWNASTGDLSYIVVNLKPGTSYQFEIFPRGPNGTEGPSQTVVGRTDC 420

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||| || || |:|||: |:|:: | || | ||||:||::||  |::  | ||: || |||
Db        421 SAVTDIRVVSVSTTEIQLEWQNTDSASGYTYHLVLESENGSIKTNSSQKWITVGGLTPGT 480

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              |||:|| |||| : |: :|  |||||||| |:|:|||    : |:| | || :||| :|
Db        481 LYNVTIFPEVDQMEGNSSSITQYTRPSNVSYIEVNTNTTVGAIQWKNLDAASASYSYSVL 540

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIF-----AQVGDGIKSLEPGRKSF 715
              | |||: || |  | |  |  |:  |||| |  |:||      :||: :    ||:| |
Db        541 ILKAGDGSNVTSRVRD--IPSVTIPGLIPGVSSEVKIFTKIRNTEVGNEV----PGQKLF 594

Qy        716 CTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTE 775
              | :|| : |  |||||||||||||| |||| |:|||| | | ||:| ||||:|: :| ||
Db        595 CMEPAQVDSLHCEVVPKEPALVLKWACPPGMNSGFELGVRSDAWDNMTHLENCTLDNDTE 654

Qy        776 YRTEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKV 835
              |||||||||||||||| |:|||||| ||::|||||||||||||||||||||||||||||
Db        655 CRTEVTYLNFSTSYNISIATLSCGKMALPTQSTCTTGITDPPPPDGSPNITSVSHNSVKV 714

Qy        836 KFSGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRS 895
              |||||||||||||||:|||||| || |||||||:|||||||||||||| ||||:|
Db        715 KFSGFEASHGPIKAYAVLLTTGEAGQPSTDVLKYTYEDFKKGASDTYVTYLIRIEEKGQS 774

Qy        896 QSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYV 955
              | ||| | ||||||:|||||||:|||||||||||||||||:: || |||:| ||||
Db        775 QGLSEALNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYV 834

Qy        956 SFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKP 1015
              ||| ||:|||||||||||||||||||| || |||||||||||||||||||||||||||
Db        835 SFSPYSEAVSLPQDPGVICGAVFGCIFGALAIVAVGGFIFWRKKRKDAKNNEVSFSQIKP 894

Qy       1016 KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVL 1075
              |||||||||||||||||||||||||||||:||| ||||||:||||||||||||
Db        895 KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYAAEIAENRGKNRYNNVL 954

Qy       1076 PYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 PYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1014

Qy       1136 IMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPL 1195
              :|||||||||||||||||||||||||||||||||:|||||||||| |||:|:||||||
Db       1015 VMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEVVLPEWTIRDFVVKNMQSSESHPL 1074

Qy       1196 RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRL 1255
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
```

```
Db       1075 RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRL 1134

Qy       1256 IYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTT 1315
              ||||||||||||||||||||||||||||||||||||||||:|:|||||||||||||||
Db       1135 IYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTT 1194

Qy       1316 AMTIYENLAPVTTFGKTNGYIA 1337
              ||||||||  |: ||| |||||
Db       1195 AMTIYENLERVSMFGKANGYIA 1216
```

RESULT 7
ADP74604
ID ADP74604 standard; peptide; 341 AA.
XX
AC ADP74604;
XX
DT 26-AUG-2004 (first entry)
XX
DE Amino acid sequence of a human DEP-1 fragment.
XX
KW human; density enhanced phosphatase-1; DEP-1; DEP-1 complex;
KW p120 catenin; adaptor protein; Gab 1; hepatocyte growth factor receptor;
KW HGF receptor; scatter factor receptor; SF receptor; Met; plakoglobin;
KW beta-catenin; signal transduction; cell growth; cell proliferation;
KW cell cycle regulation; contact inhibition; cellular differentiation;
KW cellular morphogenesis; cellular motility; cytoskeleton.
XX
OS Homo sapiens.
XX
PN WO2004048549-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038089.
XX
PR 26-NOV-2002; 2002US-0429746P.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
XX
PI Palka-Hamblin HL, Tonks NK;
XX
DR WPI; 2004-450367/42.
XX
PT New isolated complex comprises Density Enhanced Phosphatase-1 (DEP-1)
PT polypeptide, useful for manipulating biological signal transduction
PT pathways, or determining additional molecular components of the pathways.
XX
PS Claim 2; Page 98; 130pp; English.
XX
CC The present sequence represents a human density enhanced phosphatase-1
CC (DEP-1) fragment, comprising residues 997-1337 of ADP74603. The
CC specification describes DEP-1 complexes, comprising DEP-1 and DEP-1
CC substrates. DEP-1 substrate polypeptides that interact specifically with
CC the substrate-trapping mutant form of DEP-1 are isolated from human
CC breast tumour cell lines. These polypeptides include the functional
CC component of p120 catenin, the adaptor protein Gab 1, the hepatocyte
CC growth factor (HGF)/scatter factor (SF) receptor Met, a plakoglobin
CC polypeptide, and a beta-catenin polypeptide. The DEP-1 complexes are
CC useful for manipulating biological signal transduction pathways,
CC including defining therapeutic targets, or determining additional
CC molecular components of such pathways. Agents which modulate the
CC complexes are useful for therapeutic interventions which influence
CC biological processes, e.g. cell growth or proliferation including cell
CC cycle regulation and contact inhibition of cell growth, cellular
CC differentiation including altered cellular morphogenesis or motility or
CC other cellular activities characterized by alterations in cytoskeletal
CC organization and/or cellular gene expression, or cell survival including
CC cellular responses to apoptotic stimuli.
XX
SQ Sequence 341 AA;

Query Match 25.8%; Score 1811; DB 8; Length 341;
Best Local Similarity 100.0%; Pred. No. 1e-112;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        997 RKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQP 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQP 60

Qy       1057 KYAAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPL 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KYAAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPL 120

Qy       1117 PNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPE 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPE 180

Qy       1177 WTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPIL 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 WTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPIL 240

Qy       1237 VHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVL 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVL 300

Qy       1297 DIVRSQKDSKVDLIYQNTTAMTIYENLAPVTTFGKTNGYIA 1337
              ||||||||||||||||||||||||||||||||||||||||
Db        301 DIVRSQKDSKVDLIYQNTTAMTIYENLAPVTTFGKTNGYIA 341
```

RESULT 8
AAO24268
ID AAO24268 standard; protein; 1998 AA.
XX
AC AAO24268;
XX
DT 11-MAR-2004 (first entry)
XX
DE Murine vascular endothelial-protein tyrosine phosphatase.
XX
KW Vascular endothelial protein-tyrosine phosphatase; VE-PTP; mouse; human;
KW gene therapy; cytostatic; VE-cadherin; vascular endothelial-cadherin.
XX
OS Mus sp.
XX
PN WO2003084565-A2.
XX
PD 16-OCT-2003.
XX
PF 08-APR-2003; 2003WO-EP003645.
XX
PR 08-APR-2002; 2002EP-00007837.
XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI Nawroth R, Deutsch U, Vestweber D, Shima DT, Golding M;
XX
DR WPI; 2003-804251/75.
XX
PT Use of the polypeptide comprising vascular endothelial-protein tyrosine
PT phosphatase (VE-PTP) or the nucleic acid encoding the polypeptide for the
PT manufacture of an agent for monitoring or modulating VE-cadherin mediated
PT disorders.
XX
PS Example; Fig 2; 0pp; English.
XX
CC The present invention relates to a polypeptide comprising vascular
CC endothelial-protein tyrosine phosphatase (VE-PTP) or its active fragment
CC or effector, or the nucleic acid encoding the polypeptide or its
CC effector, for use in the manufacture of an agent for monitoring or
CC modulating VE-cadherin mediated processes or disorders. The polypeptide
CC comprising vascular endothelial-protein tyrosine phosphatase (VE-PTP) or
CC its active fragment or effector, or the nucleic acid encoding the
CC polypeptide or its effector, is useful for the manufacture of an agent
CC for monitoring or modulating VE-cadherin mediated processes or disorders,

```
CC   e.g., cancer. The present sequence is a protein shown in the
CC   exemplification of the invention
XX
SQ   Sequence 1998 AA;

  Query Match             17.9%;  Score 1259.5;  DB 7;  Length 1998;
  Best Local Similarity   28.8%;  Pred. No. 1.5e-74;
  Matches  378;  Conservative  216;  Mismatches  500;  Indels  217;  Gaps   45;

Qy        104 NGTDGASQKTPSSTGPSPVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTI 163
              :|   ||::    | | || |:  :: |  :: :||   :    :|:  |:    : |
Db        805 SGQYEASEQGTGRTIPEPVKDLTLLNRSTEDLHVTWSRANGDVDQYE--VQLLFNDMKVF 862

Qy        164 TVVHQPWCNITGLRPATSYVF-SITPG-----IGNETWGDPR---VIKVITEPIPVSDLR 214
                :|        :  || | | ::|||     :     || :    |: :  |  | ::
Db        863 PHIHL-------VNTATEYKFTALTPGRHYKILVLTISGDVQQSAFIEGLPVPSTVKNIH 915

Qy        215 VALTGVR-KAALSWSNGNGTASCRVLLESIGSHEELTQDSRLQVNISDLKPGVQYNINPY 273
              ::  |   :  ::|| | |                                    ::: |
Db        916 ISANGATDRLMVTWSPGGG----------------------------------DVDSY 939

Qy        274 LLQSNKTKGDPLGTEGGLDASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLV 333
              :: : :                 :  :  | | | |                   :|
Db        940 VVSAFR----------------QDEKVDSQTIPKH----------------ASEHTFH 965

Qy        334 GLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVTAVN-ISATSLTLIWKVSDNE 392
              || | :|   : |  :     |  |:  :|    |  | | |  |: |||: |: :
Db        966 RLEAGAKYRIAIVS-VSGSLRNQIDALG-QTVPASVQGVVAANAYSSNSLTVSWQKALGV 1023

Qy        393 SSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGLRSSTFYNITVCPVLGDIEGT 444
              :     | | :  |       |||||          : || :    | : :  | | :
Db       1024 AE---RYDILLLNENGLLLSNVSEPATARQHKFEDLTPGKK----YKMQILTVSGGLFSK 1076

Qy        445 PGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQMHI---TQEGAGNSRVEI-TT 500
                :  | |  |:: |:|  |:  :   |:: : |    :|     :      | ::
Db       1077 ESQAEGRTVPAAVTNLRITENSSRYLSFGWTASEGELSWYNIFLYNPDRTLQERAQVDPL 1136

Qy        501 NQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHVVYVTTTE-MWLD 559
              ||      |  |   |    ||    :|       :  ||||:|| :   :  ||: :|
Db       1137 VQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVPAAVNHLKGSHRNTTDSLWFS 1194

Qy        560 WKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGLIPG---TLYNITISPEVDHV 613
              |    |  :: | |:: : :|:   :  :| :|    |||:||   ||| :| | :: :
Db       1195 WSPASGDFDF-YELILYNPNGTKKENWKEKDVTEWRFQGLVPGRKYTLYVVTHSGDLSNK 1253

Qy        614 WGDPNSTAQYTRPSNVSNIDVSTNTTAATL----SWQNFDDAS----PTYSYCLLIEKAG 665
                ||  : || :|  ||:  : | |       | :::|      |  :  :    :
Db       1254 VTGEGRTAP-SPPSLLSFADVANTSLAITWKGPPDWTDYNDFELQWFPGDALTIFNPYSS 1312

Qy        666 NSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSL-EPGRKSFCTDPASMAS 724
              |       |  | || ||    :    ||  |: :|   | | |  : :
Db       1313 RKSEG-----------RIVYGLHPGRSYQFSVKTVSGDSWKTYSKPISGSVRTKPDKIQN 1361

Qy        725 FDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEVTYLN 784
              |       |:  |  |     |: :|          :::   |   :   | : ||
Db       1362 LHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CRKMDTQEIEFSRKLEKEKSLLN 1412

Qy        785 F-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNI------TSVSHNSV 833
                  | :||   | |  :    ::  | |  || |   |:|        :| :|:
Db       1413 IMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPQP--PPHIRVNEKDVLISKSSI 1470

Qy        834 K--VKFSGFEASHGPIKAYAVILTTG--------EAGHPSADVLKYTYDDFKKGASDTYV 883
              |  | |  ::| :| :||::          |  ||    |:| ::   :      |
Db       1471 NFTVNCSWFSDTNGAVKYFAVVVREADSMDELKPEQQHPLPSYLEYRHNASIRVYQTNY- 1529

Qy        884 TYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG---------YYNGKLEPLGSYRACVAG 934
               :  :  |   | | |    : | :| |  :||         : :|  |:|  :||  :
Db       1530 -FASKCAESPDSSSKS----FNIKLGAEMDSLGGKCDPSQQKFCDGPLKPHTAYRISIRA 1584

Qy        935 FTNITFHPQNKGLIDGAESYVSFSRYSDAVSLPQDP--GVICGAVFGCIFGALVIVTVGG 992
              || :     :: |  :   :    | : :|  ::   :|  |||  |    |     :::  |
Db       1585 FTQLF----DEDLKEFTKPLYSDTFFSMPITTESEPLFGVIEGVSAGLFLIGMLVALVAF 1640

Qy        993 FIFWRK--KRKDAKNNEVSFSQIKP----------KKSKLIRVENFEAYFKKQQADS 1037
              ||  :|    ::  :  :|  : :|             | |  |::  || :| | ||||
Db       1641 FICRQKASHSRERPSARLSIRRDRPLSVHLNLGQKGNRKTSCPIKINQFEGHFMKLQADS 1700

Qy       1038 NCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYIN 1096
              |   ::|||||| || ||    | | |||||||||:|||| |||||| |      ||||
Db       1701 NYLLSKEYEDLKDVGRSQSCDIALLPENRGKNRYNNILPYDASRVKLSNVDDDPCSDYIN 1760

Qy       1097 ANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSK 1156
              |:|:|| : ::::|||||||| |  |||:| ||:||: |:|:|:|||:|| ||: |||:
Db       1761 ASYIPGNNFRREYIATQGPLPGTKDDFWKMAWEQNVHNIVMVTQCVEKGRVKCDHYWPAD 1820

Qy       1157 Q-AQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPL-RQFHFTSWPDHGVPDTTDL 1214
              |     |||: : | || |||||||:| : : :  ::| | | ||:| |||||||:||
Db       1821 QDPLYYGDLILQMVSESVLPEWTIREFKICSEEQLDAHRLIRHFHYTVWPDHGVPETTQS 1880

Qy       1215 LINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDL 1274
              || |   ||||: :||   | :|||||||||||||:|:||:: |:::::|:|:|| |:||
Db       1881 LIQFVRTVRDYINRSPGAGPTVVHCSAGVGRTGTFVALDRILQQLDSKDSVDIYGAVHDL 1940

Qy       1275 RMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIYENLAP 1325
              |:||  ||||| |||:|:||| |::|::|        |      ||||: |
Db       1941 RLHRVHMVQTECQYVYLHQCVRDVLRAKKLRN----EQENPLFPIYENVNP 1987


RESULT 9
AAB19774
ID   AAB19774 standard; protein; 1997 AA.
XX
AC   AAB19774;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Human protein tyrosine phosphatase HPTP-beta.
XX
KW   Protein tyrosine phosphatase; HPTP-beta; human; Tie-2;
KW   receptor-type tyrosine kinase; antiangiogenic; antitumour;
KW   antimetastatic; tumour; metastasis; angiogenesis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1622. .1641
FT                   /note= "transmembrane domain"
FT   Domain          1722. .1967
FT                   /note= "catalytic domain"
XX
PN   EP1046715-A1.
XX
PD   25-OCT-2000.
XX
PF   23-APR-1999;   99EP-00108074.
XX
PR   23-APR-1999;   99EP-00108074.
XX
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI   Fachinger G,  Risau B,  Deutsch U;
XX
DR   WPI; 2000-648932/63.
DR   N-PSDB; AAA88866.
XX
PT   Monitoring or modulating Tie-2 tyrosine kinase activity, useful e.g. for
PT   regulating tumor growth, using vascular-endothelial protein tyrosine
PT   phosphatase.
XX
PS   Disclosure; Page 21-27; 60pp; English.
XX
CC   The present sequence is that of human protein tyrosine phosphatase HPTP-
CC   beta, a member of subclass III receptor type PTPs, bearing fibronectin
CC   type III-like repeats in the extracellular domain and a single catalytic
CC   domain in the cytoplasmic tail. HPTP-beta is a vascular-endothelial
CC   protein tyrosine phosphatase (VE-PTP) that specifically interacts with
```

```
CC   receptor-type tyrosine kinase Tie-2, modulating its tyrosine
CC   phosphorylation. Tie-2 is involved in angiogenetic processes, the
CC   formation of blood vessels during embryonal development, wound healing
CC   and in pathological processes such as tumour development. VE-PTPs such as
CC   HPTP-beta or its catalytic domain, nucleic acids and ligands can be used
CC   to monitor, stimulate or repress Tie-2 activity for the purpose of
CC   monitoring or modulating angiogenesis, inducing or inhibiting vascular
CC   growth or remodelling and blood vessel maturation, and inhibiting tumour
CC   growth or metastasis
XX
SQ   Sequence 1997 AA;

  Query Match             17.7%;  Score 1242;  DB 3;  Length 1997;
  Best Local Similarity   29.6%;  Pred. No. 2.2e-73;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
Db        759 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL-----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
Db        845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG-----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
Db       1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKRDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325
Db       1970 RS----EQENPLFPIYENVNP 1986
```

```
RESULT 10
AAM78821
ID   AAM78821 standard; protein; 1997 AA.
XX
AC   AAM78821;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human protein SEQ ID NO 1483.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US004098.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
PR   20-JUN-2000; 2000US-00598075.
PR   19-JUL-2000; 2000US-00620325.
PR   01-SEP-2000; 2000US-00654936.
PR   15-SEP-2000; 2000US-00663561.
PR   20-OCT-2000; 2000US-00693325.
PR   30-NOV-2000; 2000US-00728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y;
PI   Ma Y,  Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
```

```
XX
DR   WPI; 2001-476283/51.
DR   N-PSDB; AAK51954.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT   in diagnosis and gene therapy.
XX
PS   Claim 20; Page 3761-3764; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 1997 AA;

  Query Match             17.7%;  Score 1242;  DB 4;  Length 1997;
  Best Local Similarity   29.6%;  Pred. No. 2.2e-73;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
              ||   |:|   |  :  ||             | |||          |    | |   |
Db        759 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
                ::|    |    ||  |: | |     | || || :         ::||  ||
Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL-----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
                  :  :|   |  | ::           :   : | || ||      |:  : ||
Db        845 ---DVDQYEIQLLFNDMKVFPPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG-----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
              :|           | |         | || |  :          :  | | | |
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
                          |     || | :|   : |   :|:      :  ||    | |
Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
              | |  |: || : |: :   :      | | :  |      | |||           : ||
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
              :    | | :  | | :       :  | |  |:| |:|  ||  :   |:: : |
Db       1060 K----YKIQILTVSGGLFSKEAQTEERTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              :|     :|    | ::    ||      |  | |   ||    :|       :  ||||
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
              ::|  :       ||: :| :|    |  :: | |:: : :|:   :  || :|    |||
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
              :||  | : :   | |      |     :|     || :|  |::  : | |      | :
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
              ::|    :       :      |  :  ::|             | || ||   :    ||
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
              -    |:  :|   |  | |  : :  |        |:   |  |      |: :|
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
                :::   |   :   | : ||          | :||   | |  :    ::  | |  |||
Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
              |  |:|        :| :|:   |  | |  ::| :| : |::             |  ||
Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
                  |:| ::   :      |  :  :  |   | | |    :  | :| |  :||
Db       1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKCDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
                : :| |:|  :||  :  ||           | |  |     |::  |||   |||
Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
               :|  : ||: |   |  :|   | |  :  |:|    ::  :  :|  : :|
Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy        1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
                       | |  |::  || :| | |||||   ::|||:|| || :|    | | ||||
Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy        1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
               ||||||:|||| :||||| |       |||||:|:|| : ::::| |||||| |  |||:|
Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy        1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
               |||:||: |:|:|:|||:|| ||: |||: | :  |||: : | || ||||||||:| :
Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy        1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
                 : ::| | | ||:| |||||||:||  || |   ||||: :||   | :||||||||
Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy        1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
               |||||||:||:: |:::::||:|| |:|||:||  |||| |||:|:||| |::|::|
Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy        1305 SKVDLIYQNTTAMTIYENLAP 1325
                          |     ||||: |
Db       1970 RS----EQENPLFPIYENVNP 1986


RESULT 11
ADO04580
ID   ADO04580 standard; protein; 1997 AA.
XX
AC   ADO04580;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Human HPTPbeta protein.
XX
KW   Protein co-ordinate data; HPTPbeta; HPTP-beta; PTPRB; PTPbeta; PTPB;
KW   R-PTP-beta; angiogenesis mediated disorder; diabetic retinopathy;
KW   sickle cell anaemia; Paget's disease; mycobacterial infection;
KW   systemic lupus erythematosus; myopia; Crohn's disease; psoriasis;
KW   rheumatoid arthritis; tumour; acquired immune deficiency syndrome; AIDS;
KW   drug designing; therapy; human.
XX
OS   Homo sapiens.
XX
PN   US2004077065-A1.
XX
PD   22-APR-2004.
XX
PF   04-AUG-2003; 2003US-00634027.
```

XX
PR 25-SEP-2002; 2002US-0413547P.
XX
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Evdokimov AG, Pokross ME;
XX
DR WPI; 2004-374235/35.
DR N-PSDB; ADO04579.
DR GENBANK; X54131.
XX
PT Identification of compound useful for treatment of angiogenesis mediated
PT disorder, by using three-dimensional structure of HPTPbeta catalytic
PT domain, and employing structure to design, or select compound that binds
PT HPTPbeta in silico.
XX
PS Disclosure; SEQ ID NO 2; 335pp; English.
XX
CC The invention relates to the three dimensional coordinates of HPTPbeta
CC (also known as HPTP-beta, PTPRB, PTPbeta, PTPB or R-PTP-beta) protein. It
CC also relates to a method for the identification of a compound useful for
CC the treatment of an angiogenesis mediated disorder. The compounds
CC identified by this method are useful to treat diseases like diabetic
CC retinopathy, sickle cell anaemia, Paget's disease, mycobacterial
CC infections, systemic lupus erythematosus, myopia, Crohn's disease,
CC psoriasis, rheumatoid arthritis, solid or blood borne tumours and
CC acquired immune deficiency syndrome (AIDS). The invention is useful for
CC the treatment of an angiogenesis mediated disorder or disease. It is also
CC useful in drug design techniques. The present sequence is human HPTPbeta
CC protein.
XX
SQ Sequence 1997 AA;

Query Match 17.7%; Score 1242; DB 8; Length 1997;
Best Local Similarity 29.6%; Pred. No. 2.2e-73;
Matches 397; Conservative 190; Mismatches 503; Indels 251; Gaps 53;

```
Qy       123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
Db       759 FDHYEVTIKNKNNFIQTKSIP----------KSENE----------CVFVQLVPGRLY 796

Qy       183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
Db       797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy       239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
Db       845 ---DVDQYEIQLLFNDMKVFPPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy       281 KG----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
Db       902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy       314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
Db       958 ----------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy       374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy       425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy       485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy       541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy       597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy       648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy       705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy       764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy       819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy       863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
Db       1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKRDPT 1562

Qy       917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy       968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325
Db       1970 RS----EQENPLFPIYENVNP 1986
```

RESULT 12
AAE20278
ID AAE20278 standard; protein; 1997 AA.
XX
AC AAE20278;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human lung specific gene (LSG) protein #16.
XX
KW Human; lung specific gene; LSG; lung embryonic development; cytostatic;
KW lung cancer; vaccine; gene therapy; non-cancerous lung disease; tumour.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

```
FT   Region          145. .155
FT                   /note= "Antigenic epitope"
FT   Region          453. .465
FT                   /note= "Antigenic epitope"
FT   Region          650. .660
FT                   /note= "Antigenic epitope"
FT   Region          714. .735
FT                   /note= "Antigenic epitope"
FT   Region          802. .823
FT                   /note= "Antigenic epitope"
FT   Region          1040. .1062
FT                   /note= "Antigenic epitope"
FT   Region          1121. .1132
FT                   /note= "Antigenic epitope"
FT   Region          1281. .1295
FT                   /note= "Antigenic epitope"
FT   Region          1307. .1318
FT                   /note= "Antigenic epitope"
FT   Region          1334. .1347
FT                   /note= "Antigenic epitope"
FT   Region          1351. .1369
FT                   /note= "Antigenic epitope"
FT   Region          1376. .1408
FT                   /note= "Antigenic epitope"
FT   Region          1532. .1545
FT                   /note= "Antigenic epitope"
FT   Region          1644. .1655
FT                   /note= "Antigenic epitope"
FT   Region          1671. .1681
FT                   /note= "Antigenic epitope"
FT   Region          1725. .1738
FT                   /note= "Antigenic epitope"
FT   Region          1748. .1759
FT                   /note= "Antigenic epitope"
FT   Region          1804. .1825
FT                   /note= "Antigenic epitope"
FT   Region          1885. .1898
FT                   /note= "Antigenic epitope"
FT   Region          1967. .1976
FT                   /note= "Antigenic epitope"
XX
PN   WO200208278-A2.
XX
PD   31-JAN-2002.
XX
PF   20-JUL-2001; 2001WO-US022949.
XX
PR   21-JUL-2000; 2000US-0219834P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Nair M,  Chen S;
XX
DR   WPI; 2002-268964/31.
XX
PT   Novel lung specific gene useful for identifying, diagnosing, monitoring,
PT   staging, imaging and treating lung cancer and non-cancerous disease
PT   states in lung, for gene therapy, and for identifying lung tissue.
XX
PS   Claim 2; Page 185-192; 197pp; English.
XX
CC   The present invention relates to lung specific genes (LSG) and their
CC   corresponding polypeptides. LSG is useful for identifying, diagnosing,
CC   monitoring, staging, imaging and treating lung cancer and non-cancerous
CC   disease states in lung, identifying lung tissue, monitoring and modifying
CC   lung embryonic development and differentiation, in gene therapy, as
CC   hybridisation probes, to detect LSG mRNA as a marker for lung cancer, as
CC   research reagents and materials for discovery of treatments and
CC   diagnostics to human disease, to detect complementary polynucleotides,
CC   and for chromosome identification. An antibody which binds LSG is useful
CC   to detect or image localisation of LSG in a patient for detecting or
CC   diagnosing a disease or condition, for preventing the onset and treatment
CC   of lung cancer, to isolate or to identify clones expressing LSG
CC   polypeptides, to purify LSG polypeptides, and to target tumours
CC   expressing LSG. The present sequence is human LSG protein
XX
SQ   Sequence 1997 AA;

  Query Match             17.7%;  Score 1241;  DB 5;  Length 1997;
  Best Local Similarity   29.6%;  Pred. No. 2.6e-73;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
              ||   |:|   |  :  ||              | |||          |   | |  |
Db        759 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
                ::|   |    ||  |: | |     | || || :          ::||  ||
Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
                :  :|   |  | ::          :   : | || ||      |:  : ||
Db        845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG-----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
              :|           | |         | || |  :           :  | | | |
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
                          |     || | :|   : |   :|:      :  ||    | |
Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
              | |  |: || : |: :   :     | | :  |       | |||          : ||
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
              :    | | :  | | :       :  | |  |:| |:|  ||  :   |:: : |
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              :|     :|    | ::    ||      |  |  |    ||    :|       :  ||||
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
              ::|  :      ||: :| :|    |  :: | |:: : :|:   :  || :|   |||
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
              :||  | : :   | |       |     :|    || :|  |::  : | |      | :
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
              ::|   :      :      |  :  ::|              | || ||   :    ||
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
              |:  :|   |  | |  : :  |       |:   |  |      |: :|
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
              :::   |   :   | : ||          | :||   | |  :    ::  | |  |||
Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
              |   |:|      :| :|:   |  | |  ::| :| : |::            |  ||
Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
                |:| ::   :      |  :  :  |   | | |    : | :| |  :||
Db       1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKCDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
                : :| |:|  :||  :  ||          | |  |    |::  |||   |||
Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611
```

```
Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325
Db       1970 RS----EQENPLFPIYENVNP 1986
```

RESULT 13
ABR57179
ID ABR57179 standard; protein; 1997 AA.
XX
AC ABR57179;
XX
DT 02-SEP-2003 (first entry)
XX
DE Human PTPRB protein SEQ ID NO:4.
XX
KW Tec; protein tyrosine kinase; protein tyrosine phosphatase; PTP10D;
KW egg derived tyrosine phosphatase; EDTP; antidiabetic; hypotensive;
KW cardiant; antilipaemic; osteopathic; cytostatic; anorectic; obesity;
KW immunomodulator; gene therapy; metabolic disease; eating disorder;
KW body weight regulation disorder; cachexia; diabetes mellitus; cancer;
KW hypertension; coronary heart disease; hypercholesterolaemia; gallstone;
KW dyslipidaemia; osteoarthritis; sleep apnea; human; chromosome 12;
KW protein tyrosine phosphatase receptor type B precursor; PTPRB.
XX
OS Homo sapiens.
XX
PN WO2003047611-A2.
XX
PD 12-JUN-2003.
XX
PF 04-DEC-2002; 2002WO-EP013744.
XX
PR 04-DEC-2001; 2001EP-00128844.
PR 07-DEC-2001; 2001EP-00129138.
PR 02-JAN-2002; 2002EP-00000010.
XX
PA (DEVE-) DEVELOGEN ENTWICKLUNGSBIOLOGISCHE FORSCH.
XX
PI Meise M, Eulenberg K, Fritsch R, Haeder T, Broenner G;
PI Steuernagel A;
XX
DR WPI; 2003-532801/50.
DR N-PSDB; ACC79776.
XX
PT New compositions comprising tyrosine phosphatase PTP10D, protein tyrosine
PT kinase Tec or egg-derived tyrosine phosphatase genes or proteins, useful
PT for treating or preventing metabolic diseases, e.g. as obesity or
PT cachexia.
XX
PS Claim 2; Fig 8B; 83pp; English.
XX
CC The present invention describes a pharmaceutical composition comprising a
CC nucleic acid (I) protein tyrosine phosphatase PTP10D, non-receptor
CC protein tyrosine kinase Tec, egg derived tyrosine phosphatase (EDTP) gene
CC family or encoded polypeptide, fragment or variant of nucleic acid
CC molecule or polypeptide, an antibody, an aptamer or receptor recognising
CC a nucleic acid molecule of PTP10D, Tec, or EDTP gene family or encoded
CC polypeptide, and a carrier, diluent and/or adjuvant. The pharmaceutical
CC composition can have antidiabetic, hypotensive, cardiant, antilipaemic,
CC osteopathic, cytostatic, anorectic and immunomodulator activities, and
CC can be used in gene therapy. The composition is useful for the
CC manufacture of an agent for detecting and/or verifying, for treating and
CC alleviating and/or preventing a disorder, including metabolic diseases
CC such as obesity and other body weight regulation disorders, as well as
CC related disorders such as eating disorder, cachexia, diabetes mellitus,
CC hypertension, coronary heart disease, hypercholesterolaemia,
CC dyslipidaemia, osteoarthritis, gallstones, cancers (cancers of the
CC reproductive organ), sleep apnea, and other diseases, in cells, cell
CC masses, organs and/or subjects. The components of the composition may
CC also be used in controlling the function of a gene and/or gene product
CC which is influenced and/or modified by a PTP10D, Tec, or EDTP homologous
CC polypeptide, and for identifying substances capable of interacting with a
CC PTP10D, Tec or EDTP homologous polypeptide. The nucleic acid molecule of
CC PTP10D, Tec, or EDTP family or their fragments, may be used in the
CC preparation of a non-human animal which over- or under-expresses the
CC PTP10D, Tec, or EDTP gene product. The present sequence represents human
CC protein tyrosine phosphatase receptor type B precursor (PTPRB), which is
CC a human PTP10B homologous sequence. Human PTPRB is located to chromosome
CC 12
XX
SQ Sequence 1997 AA;

Query Match 17.7%; Score 1241; DB 7; Length 1997;
Best Local Similarity 29.6%; Pred. No. 2.6e-73;
Matches 397; Conservative 190; Mismatches 503; Indels 251; Gaps 53;

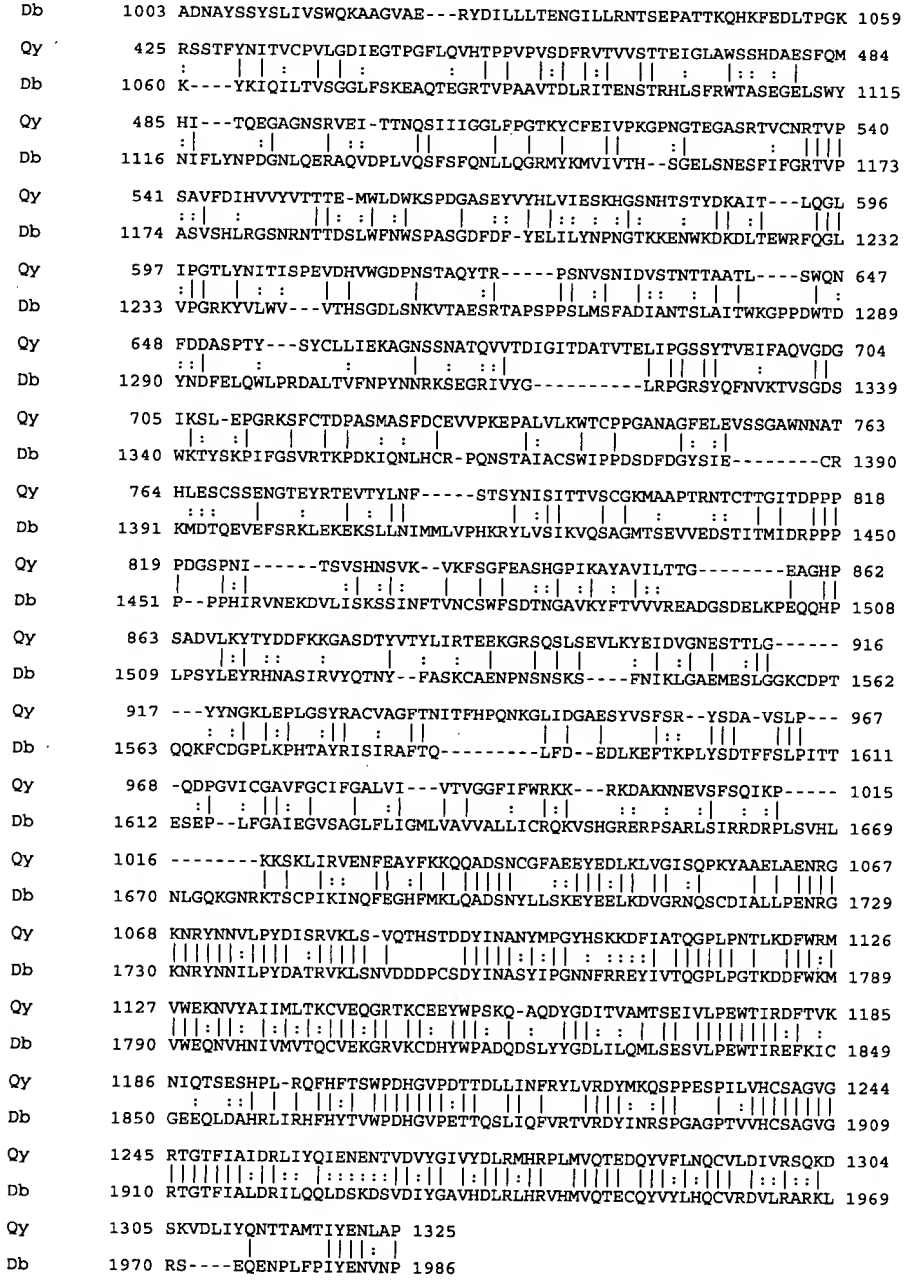
```
Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
Db        759 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
Db        845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
Db        958 ----------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232
```

```
Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
              :||  | : :   | |      |      :|     || :|  |::  : | |       | :
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
              ::|    :       :     |  :  ::|            | || ||   :    ||
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
              |:  :|   |  | |  : :   |      |:   |   |     |: :|
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
               :::   |   :   | : ||         | :||   | |  :    ::  | |  |||
Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGPEASHGPIKAYAVILTTG--------EAGHP 862
              |   |:|        :| :|:   |  | |  ::| :| : |::           |  ||
Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
              |:| ::   :     |  :  :  |   | | |    : | :| |  :||
Db       1509 LPSYLEYRHNASIRVYQTNY--PASKCAENPNSNSKS----FNIKLGAEMESLGGKCDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
               : :| |:|  :||  :  ||          | |  |    |::  |||   |||
Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
               :|  : ||: |   |  :|   |.|   :  |:|    ::  :  :|  : :|
Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
                     | |  |::  || :| | |||||   ::|||:|| || :|    | | ||||
Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
              ||||||:|||| :||||| |      |||||:|:|| : ::::| |||||| |  |||:|
Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
              |||:||: |:|:|:|||:|| ||: |||: | :  |||: : | || ||||||||:| :
Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
                :  ::| | | ||:| |||||||:||  || |   ||||: :||   | :||||||||
Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
              |||||||:||:: |:::::||:|| |:|||:|| ||||| |||:|:||| |::|::|
Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325
                    |   ||||: |
Db       1970 RS----EQENPLFPIYENVNP 1986
```

RESULT 14
ADJ70329
ID ADJ70329 standard; protein; 1997 AA.
XX
AC ADJ70329;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heat mitochondrial protein as a therapeutic target SeqID2135.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
DR WPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
PS Claim 1; SEQ ID NO 2135; 180pp; English.
XX
CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
SQ Sequence 1997 AA;

Query Match 17.7%; Score 1241; DB 7; Length 1997;
Best Local Similarity 29.6%; Pred. No. 2.6e-73;
Matches 397; Conservative 190; Mismatches 503; Indels 251; Gaps 53;

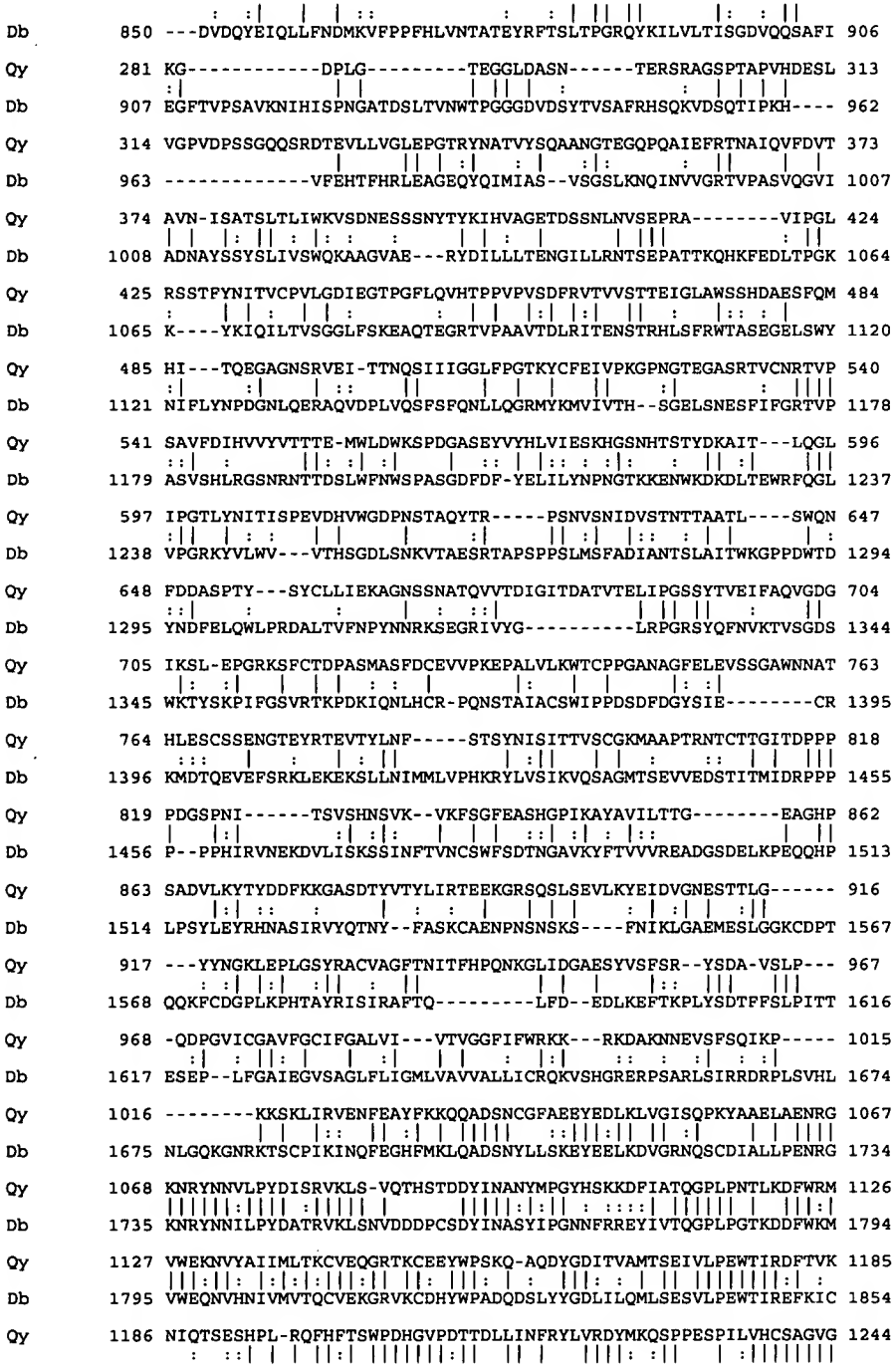
```
Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
              ||   |:|   | :  ||               | |||          |   | |   |
Db        759 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
                ::|   |    ||  |: | |     | || || :          ::||  ||
Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
                 :  :|   | | ::           :   : | || ||      |:  : ||
Db        845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
              :|              | |       | || |  :          :  | | | |
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
                          |     || | :|   : |   :|:      :  ||     |  |
Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
              | |  |: || : |: :   :     | | :  |       | |||          : ||
```

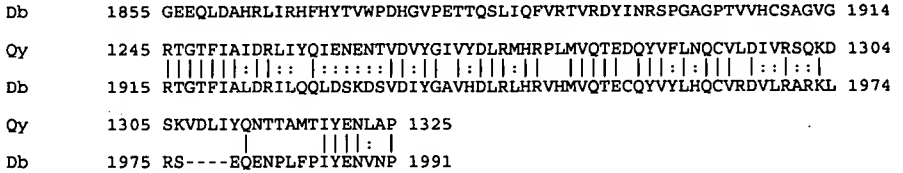


RESULT 15
AAM79805

ID AAM79805 standard; protein; 2002 AA.
XX
AC AAM79805;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3451.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK52938.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 340-341; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 2002 AA;

Query Match 17.7%; Score 1241; DB 4; Length 2002;
Best Local Similarity 29.6%; Pred. No. 2.6e-73;
Matches 397; Conservative 190; Mismatches 503; Indels 251; Gaps 53;





Search completed: May 10, 2005, 19:01:24
Job time : 210 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 10, 2005, 18:57:02 ; Search time 46 Seconds
(without alignments)
2169.691 Million cell updates/sec

Title: US-09-516-728A-4
Perfect score: 7030
Sequence: 1 MKPAAREARLPPRSPGLRWA.....TIYENLAPVTTFGKTNGYIA 1337

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6824	97.1	1337	3	US-08-854-585-2	Sequence 2, Appli
2	6824	97.1	1337	4	US-09-447-533-2	Sequence 2, Appli
3	6824	97.1	1337	5	PCT-US95-05512-2	Sequence 2, Appli
4	1242	17.7	1997	4	US-09-949-016-6275	Sequence 6275, Ap
5	1241	17.7	1903	4	US-09-949-016-8049	Sequence 8049, Ap
6	954	13.6	1711	2	US-08-342-930-2	Sequence 2, Appli
7	864	12.3	2301	4	US-09-822-871-4	Sequence 4, Appli
8	855	12.2	1188	1	US-08-201-697-4	Sequence 4, Appli
9	854	12.1	1246	4	US-09-949-016-8051	Sequence 8051, Ap
10	854	12.1	1246	4	US-09-949-016-8052	Sequence 8052, Ap
11	854	12.1	1246	4	US-09-949-016-8053	Sequence 8053, Ap
12	854	12.1	1246	4	US-09-949-016-8054	Sequence 8054, Ap
13	854	12.1	1246	4	US-09-949-016-8055	Sequence 8055, Ap
14	842	12.0	1187	1	US-08-201-697-2	Sequence 2, Appli
15	841	12.0	1274	4	US-09-949-016-8828	Sequence 8828, Ap
16	841	12.0	1274	4	US-09-949-016-8829	Sequence 8829, Ap
17	841	12.0	1274	4	US-09-949-016-8830	Sequence 8830, Ap
18	841	12.0	1274	4	US-09-949-016-8831	Sequence 8831, Ap
19	841	12.0	1274	4	US-09-949-016-8832	Sequence 8832, Ap
20	833.5	11.9	278	1	US-08-201-697-16	Sequence 16, Appl
21	826	11.7	2291	4	US-09-822-871-2	Sequence 2, Appli
22	815.5	11.6	1911	1	US-08-348-006B-5	Sequence 5, Appli
23	815.5	11.6	1911	2	US-08-800-825A-5	Sequence 5, Appli
24	815.5	11.6	1911	3	US-09-158-657-5	Sequence 5, Appli
25	815.5	11.6	1911	5	PCT-US94-10166-5	Sequence 5, Appli
26	785.5	11.2	1452	2	US-08-652-971-4	Sequence 4, Appli
27	785.5	11.2	1452	2	US-08-991-258A-4	Sequence 4, Appli
28	785.5	11.2	1452	2	US-08-769-399-4	Sequence 4, Appli
29	785.5	11.2	1452	3	US-08-991-953A-4	Sequence 4, Appli
30	778.5	11.1	254	2	US-08-685-992-14	Sequence 14, Appl
31	778.5	11.1	254	2	US-09-144-925-14	Sequence 14, Appl
32	771.5	11.0	1452	2	US-08-449-644-8	Sequence 8, Appli
33	771.5	11.0	1452	2	US-08-087-244A-8	Sequence 8, Appli
34	748	10.6	898	1	US-08-036-210-22	Sequence 22, Appl
35	748	10.6	898	2	US-08-449-609-22	Sequence 22, Appl
36	748	10.6	898	4	US-09-361-096A-22	Sequence 22, Appl
37	742.5	10.6	1501	2	US-08-447-464-3	Sequence 3, Appli
38	742.5	10.6	1501	2	US-08-716-679-3	Sequence 3, Appli
39	737.5	10.5	506	4	US-09-949-016-8833	Sequence 8833, Ap
40	737.5	10.5	506	4	US-09-949-016-8834	Sequence 8834, Ap
41	737.5	10.5	506	4	US-09-949-016-8835	Sequence 8835, Ap
42	737.5	10.5	506	4	US-09-949-016-8836	Sequence 8836, Ap
43	737.5	10.5	506	4	US-09-949-016-8837	Sequence 8837, Ap
44	737.5	10.5	506	4	US-09-949-016-8838	Sequence 8838, Ap
45	720.5	10.2	1075	4	US-09-949-016-8308	Sequence 8308, Ap

ALIGNMENTS

RESULT 1
US-08-854-585-2
; Sequence 2, Application US/08854585
; Patent No. 6114140
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,585
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,940
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-854-585-2

Query Match 97.1%; Score 6824; DB 3; Length 1337;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1302; Conservative 6; Mismatches 29; Indels 0; Gaps 0;

Qy 1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy 61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

```
Db        61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
Qy       121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
Db       121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
Qy       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
Db       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVAHGCEEGCSLSWSNGNGTASCRVLL 240
Qy       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
Db       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLAQKVAWMPAIQREAG 300
Qy       301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
Db       301 QGAPPPLCMMSPFVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
Qy       361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
Db       361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
Qy       421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
Db       421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
Qy       481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Qy       541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
Db       541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
Qy       601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
Db       601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
Qy       661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
Db       661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVELFAQVGDGIKSLEPGRKSFCTDPA 720
Qy       721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
Db       721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
Qy       781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
Db       781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
Qy       841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
Db       841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
Qy       901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
Db       901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
Qy       961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
Db       961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
Qy      1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
Db      1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
Qy      1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
Db      1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
Qy      1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
Db      1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
Qy      1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
Db      1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
Qy      1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
Db      1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
Qy      1321 ENLAPVTTFGKTNGYIA 1337
Db      1321 ENLAPVTTFGKTNGYIA 1337
```

RESULT 2
US-09-447-533-2
; Sequence 2, Application US/09447533
; Patent No. 6552169
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.
; Ostman, Arne
; TITLE OF INVENTION: DENSITY ENHANCED PROTEIN TYROSINE
; PHOSPHATASES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group PLLC
; STREET: Suite 6300, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 980104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/447,533
; FILING DATE: 23-No. 6552169-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 200125.402C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-447-533-2

Query Match 97.1%; Score 6824; DB 4; Length 1337;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1302; Conservative 6; Mismatches 29; Indels 0; Gaps 0;

```
Qy         1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
Db         1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
Qy        61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
Db        61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
Qy       121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
Db       121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
```

```
Qy       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
Db       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVAHGCEEGCSLSWSNGNGTASCRVLL 240

Qy       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
Db       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLAQKVAWMPAIQREAG 300

Qy       301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
Db       301 QGAPPPLCMMSPFVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy       361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
Db       361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy       421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
Db       421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy       481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy       541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
Db       541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy       601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
Db       601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy       661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
Db       661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVELFAQVGDGIKSLEPGRKSFCTDPA 720

Qy       721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
Db       721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy       781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
Db       781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy       841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
Db       841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy       901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
Db       901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy       961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
Db       961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy      1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
Db      1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy      1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
Db      1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy      1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
Db      1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy      1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
Db      1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy      1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
Db      1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy      1321 ENLAPVTTFGKTNGYIA 1337
Db      1321 ENLAPVTTFGKTNGYIA 1337
```

RESULT 3
PCT-US95-05512-2
; Sequence 2, Application PC/TUS9505512
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine
; TITLE OF INVENTION: Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05512
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-05512-2

Query Match 97.1%; Score 6824; DB 5; Length 1337;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1302; Conservative 6; Mismatches 29; Indels 0; Gaps 0;

```
Qy         1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
Db         1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy        61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
Db        61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy       121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
Db       121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
Db       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVAHGCEEGCSLSWSNGNGTASCRVLL 240

Qy       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
Db       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLAQKVAWMPAIQREAG 300

Qy       301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
```

```
Db        301 QGAPPPLCMMSPFVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVELFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337
```

RESULT 4
US-09-949-016-6275

```
; Sequence 6275, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6275
;   LENGTH: 1997
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6275
```

Query Match 17.7%; Score 1242; DB 4; Length 1997;
Best Local Similarity 29.6%; Pred. No. 2.7e-71;
Matches 397; Conservative 190; Mismatches 503; Indels 251; Gaps 53;

```
Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
              ||  |:|   |  :  ||                |  |||          |   |  |   |
Db        759 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
                ::|   |    ||  |: | |     | || || :         ::||  ||
Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
                :  :|   |  | ::          :   : | || ||       |:  : ||
Db        845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
              :|          | |         | || |  :            :  | | | |
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
                          |    || | :|   : |   :|:      :  ||    |  |
Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
              | |  |: || : |: :   :     | | :  |         | |||          : ||
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
              :   | | :  | |  :       :  | |  |:| |:|   ||  :   |:: : |
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              :|     :|     | ::     ||      |  |  |    ||     :|        :  ||||
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
              ::|  :       ||: :| :|     |  :: | |:: : :|:    :  || :|   |||
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
              :||  | : :   | |     |    :|     || :|  |::  : | |      | :
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
              ::|   :       :     |  :  ::|              | || ||   :    ||
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
              |:  :|   |  | |  : :  |        |:   |  |      |: :|
```

```
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818

Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862

Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916

Db       1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKRDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967

Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015

Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067

Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126

Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185

Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244

Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304

Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325

Db       1970 RS----EQENPLFPIYENVNP 1986
```

RESULT 5
US-09-949-016-8049
; Sequence 8049, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8049
;   LENGTH: 1903
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8049

Query Match 17.7%; Score 1241; DB 4; Length 1903;
Best Local Similarity 29.3%; Pred. No. 2.9e-71;
Matches 395; Conservative 198; Mismatches 507; Indels 248; Gaps 54;

```
Qy        125 IKAVSISPT----NVILTWKSNDTAASEYKYVVKH-----------KMENEKTITVVHQP 169

Db        646 VQGVSVSNSARSDYLRVSWVHATGDFDHYEVTIKNKNNFIQTKSIPKSENE--------- 696

Qy        170 WCNITGLRPATSYVFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAAL 225

Db        697 -CVFVQLVPGRLYSVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHV 749

Qy        226 SWSNGNGTASCRVLLESIGSHE--ELTQDSRL-----------QVNISDLKPGVQY---- 268

Db        750 TWSGANG---------DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILV 800

Qy        269 -NINPYLLQSNKTKG------------DPLG---------TEGGLDASN------TERSR 300

Db        801 LTISGDVQQSAFIEGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQK 860

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Db        861 VDSQTIPKH-----------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINV 901

Qy        361 EFRTNAIQVPDVTAVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419

Db        902 VGRTVPASVQGVIADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPAT 958

Qy        420 --------VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIG 471

Db        959 TKQHKFEDLTPGKK----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLS 1014

Qy        472 LAWSSHDAESFQMHI---TQEGAGNSRVEI-TTNQSIIGGLFPGTKYCFEIVPKGPNGT 527

Db       1015 FRWTASEGELSWYNIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGE 1072

Qy        528 EGASRTVCNRTVPSAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTST 586

Db       1073 LSNESFIFGRTVPASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENW 1131

Qy        587 YDKAIT---LQGLIPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNT 638

Db       1132 KDKDLTEWRFQGLVPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTS 1188

Qy        639 TAATL----SWQNFDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGS 691

Db       1189 LAITWKGPPDWTDYNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGR 1238

Qy        692 SYTVEIFAQVGDGIKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGF 750

Db       1239 SYQFNVKTVSGDSWKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGY 1297

Qy        751 ELEVSSGAWNNATHLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPT 805

Db       1298 SIE--------CRKMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVV 1349

Qy        806 RNTCTTGITDPPPPDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG 857

Db       1350 EDSTITMIDRPPPP--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREA 1407

Qy        858 --------EAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVG 909

Db       1408 DGSDELKPEQQHPLPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLG 1461

Qy        910 NESTTLG---------YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR- 959

Db       1462 AEMESLGGKCDPTQQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKP 1510

Qy        960 -YSDA-VSLP----QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNE 1007

Db       1511 LYSDTFFSLPITTESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSAR 1568

Qy       1008 VSFSQIKPKKSKL-------IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAA 1060
```

```
Db       1569 LSIRRDRPLSVHLNLGQKGPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIA 1628

Qy       1061 ELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNT 1119
Db       1629 LLPENRGKNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGT 1688

Qy       1120 LKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWT 1178
Db       1689 KDDFWKMVWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWT 1748

Qy       1179 IRDFTVKNIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILV 1237
Db       1749 IREFKICGEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVV 1808

Qy       1238 HCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLD 1297
Db       1809 HCSAGVGRTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRD 1868

Qy       1298 IVRSQKDSKVDLIYQNTTAMTIYENLAP 1325
Db       1869 VLRARKLRS----EQENPLFPIYENVNP 1892
```

RESULT 6
US-08-342-930-2
; Sequence 2, Application US/08342930
; Patent No. 5821084
; GENERAL INFORMATION:
; APPLICANT: OLMSTED, ELIZABETH A.
; APPLICANT: MAURO, LAURA J.
; APPLICANT: DAVIS, ALAN R.
; APPLICANT: DIXON, JACK E.
; TITLE OF INVENTION: OSTEOBLAST-TESTICULAR PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/342,930
; FILING DATE: 21-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KONSKI, ANTOINETTE F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: 20344-20975.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-342-930-2

Query Match 13.6%; Score 954; DB 2; Length 1711;
Best Local Similarity 25.4%; Pred. No. 9.2e-53;
Matches 368; Conservative 222; Mismatches 542; Indels 315; Gaps 64;

```
Qy          3 PAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDP---SVATVATGENGIT 59
Db        142 PARLEASWSD-APGDQDSYQLLLYHLESQTLAC---NVSVSPDTLSYSFGDLLPGTQYVL 197

Qy         60 QISSTAESFHKQN------------------GTGTPQVETNTSEDGES--------SGA 92
Db        198 EVITWAGSLHAKTSILQWTEPVPPDHLALRALGTSSLQAFWNSSEGATSFHLMLTDLLGG 257

Qy         93 NDSLRTPEQG---------SNGTDGASQKTPSSTGPSPVFDIKAVSIS----PTNVILTW 139
Db        258 TNTTAVIRQGVSTHTFLHLSPGTP-HELKICASAGPHQIWGPSATEWTYPSYPSDLVLTP 316

Qy        140 KSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSYVFSIT-PGIGNETWGDP 198
Db        317 LRNELWAS----------------------W--KAGLGARDGYVLKLSGPMESTSTLGPE 352

Qy        199 RVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLLESIGSHEELTQDSRLQVN 258
Db        353 ECNAVFPGPLPPGHYTLQLKVLAGPYDAWVEGS-----TWLAESAALPREV--------- 398

Qy        259 ISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSRA--GSPTAPVHDESLVGP 316
Db        399 -----PGARLWLD-----------------GLEASKQPGRRALLYSDDAP----GSLGN 431

Qy        317 VDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVTAVN 376
Db        432 ISVPSGA----THVIFCGLVPGAHYRVDI----ASSTGDISQSISGYTSPLPPQSLEVIS 483

Qy        377 ISATS-LTLIWKVSDNESSSNYTYKI--HVAGE----------TDSSNLNVSEPRAVIP 422
Db        484 RSSPSDLTIAWGPAPGQLEG---YKVTWHQDGSQRSPGDLVDLGPDTLSLTL---KSLVP 537

Qy        423 GLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVV-STTEIGLAW----SSH 477
Db        538 G----SCYTVSAWAWAGNLSSDSQKIHSCTRPAPPTNLSLGFAHQPAALKASWYHPPGGR 593

Qy        478 DAESFQMH-----------ITQEGAGNSRVEITTNQSIIIGGLFPGTKYC-FEIVPKGP 524
Db        594 DAFHLRLYRLRPLTLESEKVLPREAQNFSWAQLTAG-------------CEFQVQLSTL 639

Qy        525 NGTE-GASRTVCNRTVPSAVFDIHVVYVTTTEMWLDWKS-PDGASEYVYHLVIES-KHGS 581
Db        640 WGSERSSSANATGWTPPSAPTLVNVTSDAPTQLQVSWAHVPGGRSRYQVTLYQESTRTAT 699

Qy        582 NHTSTYDKAITLQGLIPGTLYNITISPEVDHVWGDPNSTAQYTRPSNVSN---------- 631
Db        700 SIMGPKEDGTSFLGLTPGTKYKVEVIS-----WAGP----LYTAAANVSAWTYPLIPNEL 750

Qy        632 -IDVSTNTTAATLSWQNFDDASPT-YSYC-LLIEKAGNSSNATQVVTDIGITDATVTELI 688
Db        751 LVSMQAGSAVVNLAW----PSGPLGQGACHAQLSDAGHLS--WEQPLKLGQELFMLRDLT 804

Qy        689 PGSSYTVEIFAQVGD-------GIKSLEPGRKSFCTDPASMASFDCEVVPKEPALVLKWT 741
Db        805 PGHTISMSVRCRAGPLQASTHLVVLSVEPG-------PVE----DVLCHPEATYLALNWT 853

Qy        742 CP--------------PGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEVTYLNFS 786
Db        854 MPAGDVDVCLVVVERLVPGGGTHFVFQVNTS--GDALLLPNLMPT--TSYRLSLTVLGRN 909

Qy        787 TSYNISITTVSCGKMAAPTRNTCTTGITDPPPPD--GSPNITSVSHNSVKVKFSGFEASH 844
Db        910 SRWSRAVSLV-----------CSTSAEAWHPPELAEPPQVELGTGMGVTVMRGMFGKDD 957

Qy        845 GPIKAYAVILTTG-EAGHPSADVLKYT-YDDFKKGASDTYVTYLIRT----EEKGRSQSL 898
Db        958 GQIQWYGIIATINMTLAQPSREAINYTWYDHYYRGC-ESFLALLFPNPFYPEPWAGPRS- 1015

Qy        899 SEVLKYEIDVGNE--STTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVS 956
Db       1016 -----WTVPVGTEDCDNTQEICNGRLKSGFQYRFSVVAFSR----------LNTPETILA 1060

Qy        957 FSRYSD---AVSLPQDP-GVICGAVFGCIFGALVIVTVGGFIFWR-----KKRKDAKNNE 1007
Db       1061 FSAFSEPRASISLAIIPLTVMLGAVVGSI---VIVCAVLCLLRWRCLKGPRSEKDGFSKE 1117
```

```
Qy       1008 VSFSQIKP----KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELA 1063
Db       1118 -----LMPYNLWRTHRPIPIHSFRQSYEAKSAHAHQTFFQEFEELKEVGKDQPRLEAEHP 1172

Qy       1064 ENRGKNRYNNVLPYDISRVKLSV---QTHSTDDYINANYMPGYHSKKDFIATQGPLPNTL 1120
Db       1173 DNIIKNRYPHVLPYDHSRVRLTQLPGEPHS--DYINANFIPGYSHTQEIIATQGPLKKTL 1230

Qy       1121 KDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQ-DYGDITVAMTSEIVLPEWTI 1179
Db       1231 EDFWRLVWEQQVHVIIMLTVGMENGRVLCEHYWPANSTPVTHGHITIHLLAEEPEDEWTR 1290

Qy       1180 RDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHC 1239
Db       1291 REFQLQHGTEQKQRRVKQLQFTTWPDHSVPEAPSSLLAFVELVQEQVQATQGKGPILVHC 1350

Qy       1240 SAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLD-I 1298
Db       1351 SAGVGRTGTFVALLRLLRQLEEEKVADVFNTVYILRLHRPLMIQTLSQYIFLHSCLLNKI 1410

Qy       1299 VRSQKDS 1305
Db       1411 LEGPPDS 1417


RESULT 7
US-09-822-871-4
; Sequence 4, Application US/09822871
; Patent No. 6723547
; GENERAL INFORMATION:
;  APPLICANT: WEBSTER, Marion et al
;  TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL001219
;  CURRENT APPLICATION NUMBER: US/09/822,871
;  CURRENT FILING DATE:  2002-12-02
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 2301
;   TYPE: PRT
;   ORGANISM: Rat
US-09-822-871-4

  Query Match           12.3%;  Score 864;  DB 4;  Length 2301;
  Best Local Similarity  23.5%;  Pred. No. 9.6e-47;
  Matches  405;  Conservative  216;  Mismatches  533;  Indels  568;  Gaps   77;

Qy         87 GESSGANDS---LRTPEQGSNGTDGASQKTPSSTGPSPVFDIKAVSISPTNVILTW---- 139
Db        647 GESSLSEENDIFVRTPED----------EPES---SPQ-DVQVTGVSPSELRLKWSPPE 691

Qy        140 KSNDTAASEYKYVVKHK----MENEKTITVVHQPWCNITGLRPATSYVFSITP----GIG 191
Db        692 KPNGIIIA-YEVLYQNADTLFVKNTSTTDII------ISDLKPYTLYNISIRSYTRLGHG 744

Qy        192 NETWG--DPRVIKVITEPIP--VSDLRVALTGVRKAALSWSNGNG-TASCRVLLESIGSH 246
Db        745 NQSSSLLSVRTSETVPDSAPENITYKNISSGEIEISFLPPRSPNGIIQKYTIYLKRSNSH 804

Qy        247 EELT-QDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSRAGSPT 305
Db        805 EARTINTTSLTQTIGGLKKYTHYVIE---VSASTLKGE--GIRSRPISILTEEDAPDSPP 859

Qy        306 APVHDESLVG---------PVDP--------------SSGQQSRDTEVLLV--GLEPGTR 340
Db        860 QNFSVKQLSGVTVMLSWQPPLEPNGIILYYTVYVWDKSSLRAINATEASLVLSDLDYNVD 919

Qy        341 YNATVYSQAANGT-EGQPQAIEFRTNAIQVF----DVTAVNISATSLTLIW--------- 386
Db        920 YGACVTASTRFGDGNARSSIINFRTPEGEPSDPPNDVHYVNLSSSSIILFWTPPVKPNGI 979

Qy        387 ------------------------KVSDNESSS------------------------- 395
Db        980 IQYYSVYYQNTSGTFVQNFTLLQVTKESDNVTVSARIYRLAIFSYYTFWLTASTSVGNGN 1039

Qy        396 --------------------NYTYK------IHVAGETDSS-----------NLNVSEPR 418
Db       1040 KSSDIIHVYTDQDIPEGPVGNLTFESISSTAIHVSWEPPSQPNGLVFYYLSLNLQQSPPR 1099

Qy        419 AVIP------------GLRSSTFYNITVCPVLGDIEGTPGF-----LQVH---------T 452
Db       1100 HMIPPLVTYENSIDFDDLEKYTDYIPKITP-----STEKGFSETYTTQLHIKTEEDVPDT 1154

Qy        453 PPVPVSDFRVTVVSTTEIGLAWS---SHDAESFQMHITQEGAGNSRVEITTNQSIIIGGL 509
Db       1155 PPI-INTFK--NLSSTSILLSWDPPLKPNGAILGYHLTLQGPHANHTFVTSGNHIVLEEL 1211

Qy        510 FPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVF----DIHVVYVTTTEMWLDWK-SP- 563
Db       1212 SPFTLYSFFAAARTMKGL-GPSSILFFYTDESAPLAPPQNLTLINYTSDFVWLTWSPSPL 1270

Qy        564 DGASEYVYHLVIESKHGSNHTSTYDKAIT-------LQGLIPGTLYNITISPEVDHVWGD 616
Db       1271 PGGIVKVYSFKI---HEHETDTVFYKNISGLQTDAKLEGLEPVSTYSVSVSAFTKVGNGN 1327

Qy        617 PNS-----TAQYTRPSNVSNID-VSTNTTAATLSWQNFDDASPTYSYCL--LIEKAGNSS 668
Db       1328 QYSNVVEFTTQESVPEAVRNIECVARDWQSVSVRW---DPPRKTNGIIIHYMITVGGNS- 1383

Qy        669 NATQVVTDIGITDA--TVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMASFD 726
Db       1384 ------TKVSPRDPTYTFTKLLPNTSYVFEVRASTSAG-----EGNESRC-DISTLP--- 1428

Qy        727 CEVVPKEPALV-----------LKWTCP------------------------------ 743
Db       1429 -ETVPSAPTNVAFSNVQSTSATLTWTKPDTIFGYFQNYKITTQLRAQKCREWEPEECIEH 1487

Qy        744 ------------------------------------------------PG--------- 745
Db       1488 QKDQYLYEANQTEETVHGLKKFRWYRFQVAASTNVGYSNASEWISTQTLPGPPDGPPENV 1547

Qy        746 ---ANAGPELEVSSGAWN----------------------NATHLESCSSENGTEYRTE 779
Db       1548 HVVATSPFGINIS---WSEPAVITGPTFYLIDVKSVDDDDFNISFLKS-NEENKT---TE 1600

Qy        780 VTYLNFSTSYNISIT--------TVSCGKMAAPTRNTCTTGITDPPP------------- 818
Db       1601 INNLEVFTRYSVVITAFVGNVSRAYTDGKSSAEVIITTLESVPKDPPNNMTFQKIPDEVT 1660

Qy        819 -------PDGSPN------------------------ITSVSHNSVKVKFSGFEASHG- 845
Db       1661 KFQLTFLPPSQPNGNIRVYQALVYREDDPTAVQIHNFSIIQKTDTSIIAMLEGLKGGHTY 1720

Qy        846 PIKAYAVILTTGEAGHPSADVLKYTYD---------------DFKKGASDTYVTYLIR-- 888
Db       1721 NISVYAI----NSAGAGPKVQMRITMDIKAPARPKSKPIPIRDATGKLLVTSTTITIRMP 1776

Qy        889 -----------------TEEKGRSQSLSEVLKYEIDVGNES----TTLGYYN-----GKL 922
Db       1777 ICYYNDDHGPIRNVQVLVAETGAQQD-GNVTKWYDAYFNKARPYFTNEGFPNPPCIEGKT 1835

Qy        923 EPLGSYRACVAGFTNITFHPQNK-----GLIDGAESYV-----------SFSRYSDAV- 964
Db       1836 KFSGNEEIYVIGADNACMIPGNEEKICNGPLKPKKQYLFKFRATNVMGQFTDSEYSDPIK 1895

Qy        965 ----SLPQDPGVICGAVFGCIFGALVIVT-VGGFIFWRKKRK--------DAKNNEVSF- 1010
Db       1896 TLGEGLSERTVEIILSVTLCILSIILLGTAIFAFVRIRQKQKEGGTYSPRDAEIIDTKFK 1955

Qy       1011 -------SQIKPKKSKLIRV------------ENFEAYFKKQQADSNCGFAEEYEDLKLV 1051
Db       1956 LDQLITVADLELKDERLTRLLSYRKSIKPISKKSFLQHVEELCTNSNLKFQEEFSEL--- 2012
```

```
Qy       1052 GISQPKYA-------AELAENRGKNRYNNVLPYDISRVKLSVQTH-STDDYINANYMPGY 1103
Db       2013 ----PKFLQDLSSTDADLPWNRAKNRFPNIKPYNNNRVKLIADVSLPGSDYINASYVSGY 2068

Qy       1104 HSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP--SKQAQDY 1161
Db       2069 LCPNEFIATQGPLPGTVGDFWRMVWETRTKTLVMLTQCFEKGRIRCHQYWPEDNKPVTVF 2128

Qy       1162 GDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYL 1221
Db       2129 GDIVITKLMEDIQIDWTIRDLKIE--RHGDCMTVRQCNFTGWPEHGVPENTTPLIHFVKL 2186

Qy       1222 VRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLM 1281
Db       2187 VR--TSRAHDTTPMVVHCSAGVGRTGVFIALDHLTQHINNHDFVDIYGLVAELRSERMCM 2244

Qy       1282 VQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIYENL 1323
Db       2245 VQNLAQYIFLHQCILDLL-SNKGGHQPVCFVNYSTLQKMDSL 2285
```

```
RESULT 8
US-08-201-697-4
; Sequence 4, Application US/08201697
; Patent No. 5705623
;  GENERAL INFORMATION:
;    APPLICANT:  Wiggins, Roger C.
;    APPLICANT:  Thomas, Peedikayil E.
;    TITLE OF INVENTION:  Mammalian Glomerular Epithelial Protein
;    TITLE OF INVENTION:  1
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/201,697
;      FILING DATE:  25-FEB-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryb A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UM 9783
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1188 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-201-697-4

  Query Match             12.2%;  Score 855;  DB 1;  Length 1188;
  Best Local Similarity   28.0%;  Pred. No. 1.3e-46;
  Matches  279;  Conservative  159;  Mismatches  345;  Indels  214;  Gaps   42;

Qy        452 TPPVPVS-------DFRVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVEIT 499
Db        271 TPEIPSGNISSGWPDFNSSDYETTSQPYWWDSASAAPESEDEFVSVLPMEYENNSTLSET 330

Qy        500 TNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHV---------- 548
Db        331 EKSTSGSFSFFPVQMILTWLPPKPPTAFDG-------------FHIHIEREEKFTEYLM 376
```

```
Qy        549 VYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISP 608
Db        377 VDEEAHEFVAELKEP---GKYKLSVTTFSSSGSCETRKSQSAKSL---------SFYISP 424

Qy        609 EVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCLLIEKA 664
Db        425 SGE--WIEELT----EKPQHVS-VHVLSSTT-ALMSWTSSQENYNSTIVSV-VSLTCQKQ 475

Qy        665 GNSSN-ATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMA 723
Db        476 KESQRLEKQYCTQVNSSKPIIENLVPGAQYQVVIYILRKGPLI-----GPPS---DPVTFA 527

Qy        724 SF-----DCEVVPKEP-ALVLKWTCP-PGANAGFELEVSSGAWNNATHLESCSSENGTEY 776
Db        528 IVPTGIKDLMLYPLGPTAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSEWTTYY 581

Qy        777 RTEVTYLNFSTS----------YNISITTVSCGKMAAPTRNTCTTG----ITDPPPPDG 821
Db        582 EIAAT-VSLTASVRIANLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAPV---- 633

Qy        822 SPNITSVSHNSVKVKFS------GFEASHGPIKAYAVI----------LTTGEAGH 861
Db        634 APEITSVEYFNSLLYISWTYGDDTTDLSHSRMLHWMVVAEGKKKIKKSVTRNVMTAILSL 693

Qy        862 PSADVLKYTYDDFKKGASDTYVTYLIRTE---------------------------- 890
Db        694 PPGDIYNLSVTACTERGSNTSMLRLVKLEPAPPKSLFAVNKTQTSVTLLWVEEGVADFFE 753

Qy        891 ----EKGRSQSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKG 946
Db        754 VFCQQVGSSQKTK--LQEPVAVSSHVVTI----SSLLPATAYNCSVTSFSH--------- 798

Qy        947 LIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKK-----RK 1001
Db        799 ---DSPSVPTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQMARE 854

Qy       1002 DAKNNEVSFSQIKP-------------KKSKL---IRVENFEAYFKKQQADSNCGFAEEY 1045
Db        855 CGAGTFVNFASLERDGKLPYNWSKNGLKKRKLTNPVQLDDFDAYIKDMAKDSDYKFSLQF 914

Qy       1046 EDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKL-SVQTHSTDDYINANYMPGYH 1104
Db        915 EELKLIGLDIPHFAADLPLNRCKNRYTNILPYDFSRVRLVSMNEEEGADYINANYIPGYN 974

Qy       1105 SKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQDYGD 1163
Db        975 SPQEYIATQGPLPETRNDFWKMVLQQKSQIIVMLTQCNEKRRVKCDHYWPFTEEPIAYGD 1034

Qy       1164 ITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVP--DTTDLLINFRYL 1221
Db       1035 ITVEMISEEEQDDWACRHFRIN--YADEMQDVMHFNYTAWPDHGVPTANAAESILQFVHM 1092

Qy       1222 VRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLM 1281
Db       1093 VRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRDHEFVDILGLVSEMRSYRMSM 1150

Qy       1282 VQTEDQYVFLNQCVLDIVRSQKDSKV---DLIYQNTT 1315
Db       1151 VQTEEQYIFIHQCV-QLMWMKKKQQICISDVIYENVS 1186
```

```
RESULT 9
US-09-949-016-8051
; Sequence 8051, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
```

```
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8051
;   LENGTH: 1246
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8051

  Query Match            12.1%;  Score 854;  DB 4;  Length 1246;
  Best Local Similarity  28.0%;  Pred. No. 1.7e-46;
  Matches  279;  Conservative  158;  Mismatches  347;  Indels  212;  Gaps   41;

Qy       452 TPPVPVS-------DFRVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVEIT 499
Db       329 TPEIPSGNISSGWPDFNSSDYETTSQPYWWDSASAAPESEDEFVSVLPMEYENNSTLSET 388

Qy       500 TNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHV----------- 548
Db       389 EKSTSGSFSFFPVQMILTWLPPKPPTAFDG-------------FHIHIEREENFTEYLM 434

Qy       549 VYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISP 608
Db       435 VDEEAHEFVAELKEP---GKYKLSVTTFSSSGSCETRKSQSAKSL---------SFYISP 482

Qy       609 EVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCLLIEKA 664
Db       483 SGE--WIEELT----EKPQHVS-VHVLSSTT-ALMSWTSSQENYNSTIVSV-VSLTCQKQ 533

Qy       665 GNSSN-ATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMA 723
Db       534 KESQRLEKQYCTQVNSSKPIIENLVPGAQYQVVIYLRKGPLI-----GPPS---DPVTFA 585

Qy       724 SF-----DCEVVPKEP-ALVLKWTCP-PGANAGFELEVSSGAWNNATHLESCSSENGTEY 776
Db       586 IVPTGIKDLMLYPLGPTAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSEWTTYY 639

Qy       777 RTEVTYLNFSTS-----------YNISITTVSCGKMAAPTRNTCTTG----ITDPPPPDG 821
Db       640 EIAAT-VSLTASVRIANLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAPV---- 691

Qy       822 SPNITSVSHNSVKVKFS------GFEASHGPIKAYAVI--------------LTTGEAGH 861
Db       692 APEITSVEYFNSLLYISWTYGDDTTDLSHSRMLHWMVVAEGKKKIKKSVTRNVMTAILSL 751

Qy       862 PSADVLKYTYDDFKKGASDTYVTYLIRTE------------------------------- 890
Db       752 PPGDIYNLSVTACTERGSNTSMLRLVKLEPAPPKSLPAVNKTQTSVTLLWVEEGVADFFE 811

Qy       891 ----EKGRSQSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKG 946
Db       812 VFCQQVGSSQKTK--LQEPVAVSSHVVTI----SSLLPATAYNCSVTSFSH--------- 856

Qy       947 LIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKK-----RK 1001
Db       857 ---DSPSVPTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQMARE 912

Qy      1002 DAKNNEVSFSQIKP-------------KKSKL---IRVENFEAYFKKQQADSNCGFAEEY 1045
Db       913 CGAGTFVNFASLERDGKLPYNWSKNGLKKRKLTNPVQLDDFDAYIKDMAKDSDYKFSLQF 972

Qy      1046 EDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKL-SVQTHSTDDYINANYMPGYH 1104
Db       973 EELKLIGLDIPHFAADLPLNRCKNRYTNILPYDFSRVRLVSMNEEEGADYINANYIPGYN 1032

Qy      1105 SKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQDYGD 1163
Db       1033 SPQEYIATQGPLPETRNDFWKMVLQQKSQIIVMLTQCNEKRRVKCDHYWPFTEEPIAYGD 1092

Qy      1164 ITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVP--DTTDLLINFRYL 1221
Db       1093 ITVEMISEEEQDDWACRHFRIN--YADEMQDVMHFNYTAWPDHGVPTANAAESILQFVHM 1150

Qy      1222 VRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLM 1281
Db       1151 VRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRDHEFVDILGLVSEMRSYRMSM 1208

Qy      1282 VQTEDQYVFLNQCV--LDIVRSQKDSKVDLIYQNTT 1315
Db       1209 VQTEEQYIFIHQCVQLMWMKKKQQFCISDVIYENVS 1244

RESULT 10
US-09-949-016-8052
; Sequence 8052, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8052
;   LENGTH: 1246
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8052

  Query Match            12.1%;  Score 854;  DB 4;  Length 1246;
  Best Local Similarity  28.0%;  Pred. No. 1.7e-46;
  Matches  279;  Conservative  158;  Mismatches  347;  Indels  212;  Gaps   41;

Qy       452 TPPVPVS-------DFRVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVEIT 499
Db       329 TPEIPSGNISSGWPDFNSSDYETTSQPYWWDSASAAPESEDEFVSVLPMEYENNSTLSET 388

Qy       500 TNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHV----------- 548
Db       389 EKSTSGSFSFFPVQMILTWLPPKPPTAFDG-------------FHIHIEREENFTEYLM 434

Qy       549 VYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISP 608
Db       435 VDEEAHEFVAELKEP---GKYKLSVTTFSSSGSCETRKSQSAKSL---------SFYISP 482

Qy       609 EVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCLLIEKA 664
Db       483 SGE--WIEELT----EKPQHVS-VHVLSSTT-ALMSWTSSQENYNSTIVSV-VSLTCQKQ 533

Qy       665 GNSSN-ATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMA 723
Db       534 KESQRLEKQYCTQVNSSKPIIENLVPGAQYQVVIYLRKGPLI-----GPPS---DPVTFA 585

Qy       724 SF-----DCEVVPKEP-ALVLKWTCP-PGANAGFELEVSSGAWNNATHLESCSSENGTEY 776
Db       586 IVPTGIKDLMLYPLGPTAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSEWTTYY 639

Qy       777 RTEVTYLNFSTS-----------YNISITTVSCGKMAAPTRNTCTTG----ITDPPPPDG 821
Db       640 EIAAT-VSLTASVRIANLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAPV---- 691

Qy       822 SPNITSVSHNSVKVKFS------GFEASHGPIKAYAVI--------------LTTGEAGH 861
```

```
Db        692 APEITSVEYFNSLLYISWTYGDDTTDLSHSRMLHWMVVAEGKKKIKKSVTRNVMTAILSL 751

Qy        862 PSADVLKYTYDDFKKGASDTYVTYLIRTE------------------------------ 890
              |  |:   :      :  |:| :  |:: |
Db        752 PPGDIYNLSVTACTERGSNTSMLRLVKLEPAPPKSLFAVNKTQTSVTLLWVEEGVADFFE 811

Qy        891 ----EKGRSQSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKG 946
                  : | ||      |:  : | :   |:       | |  :|    |  |::
Db        812 VFCQQVGSSQKTK--LQEPVAVSSHVVTI----SSLLPATAYNCSVTSFSH--------- 856

Qy        947 LIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKK-----RK 1001
                  : |  :|    |  |:    :| |:  :|    :    |: : :    |  |||      |:
Db        857 ---DSPSVPTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQMARE 912

Qy       1002 DAKNNEVSFSQIKP-------------KKSKL---IRVENFEAYFKKQQADSNCGFAEEY 1045
                |:|: ::                  || ||   :::::|:|| |    ||:  |: ::
Db        913 CGAGTFVNFASLERDGKLPYNWSKNGLKKRKLTNPVQLDDFDAYIKDMAKDSDYKFSLQF 972

Qy       1046 EDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKL-SVQTHSTDDYINANYMPGYH 1104
              |:|||:|:  | :||:|  || |||| |:|||| |||:| |:      |||||||:|||:
Db        973 EELKLIGLDIPHFAADLPLNRCKNRYTNILPYDFSRVRLVSMNEEEGADYINANYIPGYN 1032

Qy       1105 SKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQDYGD 1163
              | :::|||||||| |  |||:|| ::    |:|||:| |: | ||: ||| :::   |||
Db       1033 SPQEYIATQGPLPETRNDFWKMVLQQKSQIIVMLTQCNEKRRVKCDHYWPFTEEPIAYGD 1092

Qy       1164 ITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVP--DTTDLLINFRYL 1221
              |||  | ||     :|  | | :      |   :  |::|:|||||||  :  : :: | ::
Db       1093 ITVEMISEEEQDDWACRHFRIN--YADEMQDVMHFNYTAWPDHGVPTANAAESILQFVHM 1150

Qy       1222 VRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLM 1281
              ||     :|  : |:::|||||||||||||||:|||:  | :   ||: |:| ::| :|  |
Db       1151 VRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRDHEFVDILGLVSEMRSYRMSM 1208

Qy       1282 VQTEDQYVFLNQCV--LDIVRSQKDSKVDLIYQNTT 1315
              ||||:||:|::|||  : : : |:    |:||:| :
Db       1209 VQTEEQYIFIHQCVQLMWMKKKQQFCISDVIYENVS 1244


RESULT 11
US-09-949-016-8053
; Sequence 8053, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8053
;   LENGTH: 1246
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8053

  Query Match           12.1%;  Score 854;  DB 4;  Length 1246;
  Best Local Similarity  28.0%;  Pred. No. 1.7e-46;
  Matches  279;  Conservative  158;  Mismatches  347;  Indels  212;  Gaps   41;

Qy        452 TPPVPVS-------DFRVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVEIT 499
              || :|       ||  :   ||     | |  |     :  |   :  |   || :  |
Db        329 TPEIPSGNISSGWPDFNSSDYETTSQPYWWDSASAAPESEDEFVSVLPMEYENNSTLSET 388

Qy        500 TNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHV----------- 548
                 :      ||        : || |   :|              | ||:
Db        389 EKSTSGSFSFFPVQMILTWLPPKPPTAFDG-------------FHIHIEREENFTEYLM 434

Qy        549 VYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISP 608
              |      |   : | |     :|   :    |. || |      | :|         : |||
Db        435 VDEEAHEFVAELKEP---GKYKLSVTTFSSSGSCETRKSQSAKSL---------SFYISP 482

Qy        609 EVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCLLIEKA 664
                 : | :  :    :| :|| : | ::|| | :||    :|::    :    |  :|
Db        483 SGE--WIEELT----EKPQHVS-VHVLSSTT-ALMSWTSSQENYNSTIVSV-VSLTCQKQ 533

Qy        665 GNSSN-ATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMA 723
                 |      |  | :  :    :  |:||: | | |: : |  |      |  |   || : |
Db        534 KESQRLEKQYCTQVNSSKPIIENLVPGAQYQVVIYILRKGPLI-----GPPS---DPVTFA 585

Qy        724 SF-----DCEVVPKEP-ALVLKWTCP-PGANAGFELEVSSGAWNNATHLESCSSENGTEY 776
                    |  : |  | |:|| || |  |    : :|:    :| ||      :|| | |
Db        586 IVPTGIKDLMLYPLGPTAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSEWTTYY 639

Qy        777 RTEVTYLNFSTS-----------YNISITTVSCGKMAAPTRNTCTTG----ITDPPPPDG 821
                  | :: : |            ||  :| |: |    |  : | :     || |
Db        640 EIAAT-VSLTASVRIANLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAPV---- 691

Qy        822 SPNITSVSHNSVKVKFS------GFEASHGPIKAYAVI--------------LTTGEAGH 861
                :| |||| : :  :  |          : ||  : .: |:                : |
Db        692 APEITSVEYFNSLLYISWTYGDDTTDLSHSRMLHWMVVAEGKKKIKKSVTRNVMTAILSL 751

Qy        862 PSADVLKYTYDDFKKGASDTYVTYLIRTE------------------------------ 890
              |  |:   :      :  |:| :  |:: |
Db        752 PPGDIYNLSVTACTERGSNTSMLRLVKLEPAPPKSLFAVNKTQTSVTLLWVEEGVADFFE 811

Qy        891 ----EKGRSQSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKG 946
                  : | ||      |:  : | :   |:       | |  :|    |  |::
Db        812 VFCQQVGSSQKTK--LQEPVAVSSHVVTI----SSLLPATAYNCSVTSFSH--------- 856

Qy        947 LIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKK-----RK 1001
                  : |  :|    |  |:    :| |:  :|    :    |: : :    |  |||      |:
Db        857 ---DSPSVPTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQMARE 912

Qy       1002 DAKNNEVSFSQIKP-------------KKSKL---IRVENFEAYFKKQQADSNCGFAEEY 1045
                |:|: ::                  || ||   :::::|:|| |    ||:  |: ::
Db        913 CGAGTFVNFASLERDGKLPYNWSKNGLKKRKLTNPVQLDDFDAYIKDMAKDSDYKFSLQF 972

Qy       1046 EDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKL-SVQTHSTDDYINANYMPGYH 1104
              |:|||:|:  | :||:|  || |||| |:|||| |||:| |:      |||||||:|||:
Db        973 EELKLIGLDIPHFAADLPLNRCKNRYTNILPYDFSRVRLVSMNEEEGADYINANYIPGYN 1032

Qy       1105 SKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQDYGD 1163
              | :::|||||||| |  |||:|| ::    |:|||:| |: | ||: ||| :::   |||
Db       1033 SPQEYIATQGPLPETRNDFWKMVLQQKSQIIVMLTQCNEKRRVKCDHYWPFTEEPIAYGD 1092

Qy       1164 ITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVP--DTTDLLINFRYL 1221
              |||  | ||     :|  | | :      |   :  |::|:|||||||  :  : :: | ::
Db       1093 ITVEMISEEEQDDWACRHFRIN--YADEMQDVMHFNYTAWPDHGVPTANAAESILQFVHM 1150

Qy       1222 VRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLM 1281
              ||     :|  : |:::|||||||||||||||:|||:  | :   ||: |:| ::| :|  |
Db       1151 VRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRDHEFVDILGLVSEMRSYRMSM 1208

Qy       1282 VQTEDQYVFLNQCV--LDIVRSQKDSKVDLIYQNTT 1315
              ||||:||:|::|||  : : : |:    |:||:| :
Db       1209 VQTEEQYIFIHQCVQLMWMKKKQQFCISDVIYENVS 1244


RESULT 12
US-09-949-016-8054
; Sequence 8054, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
```

```
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8054
;   LENGTH: 1246
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8054

  Query Match             12.1%;  Score 854;  DB 4;  Length 1246;
  Best Local Similarity   28.0%;  Pred. No. 1.7e-46;
  Matches  279;  Conservative  158;  Mismatches  347;  Indels  212;  Gaps   41;

Qy        452 TPPVPVS-------DFRVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVEIT 499
              || :|         ||  :   ||      | |  |      :  |   :  |  || : |
Db        329 TPEIPSGNISSGWPDFNSSDYETTSQPYWWDSASAAPESEDEFVSVLPMEYENNSTLSET 388

Qy        500 TNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHV----------- 548
                  :       ||        : || |   :|                | ||:
Db        389 EKSTSGSFSFFPVQMILTWLPPKPPTAFDG-------------FHIHIEREENFTEYLM 434

Qy        549 VYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISP 608
              |      |   : | |     :|    :     | ||     | :|          :  |||
Db        435 VDEEAHEFVAELKEP---GKYKLSVTTFSSSGSCETRKSQSAKSL---------SFYISP 482

Qy        609 EVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCLLIEKA 664
                 :  | :  :      :| :|| : | ::|| | :||   :|::    :    |  :|
Db        483 SGE--WIEELT----EKPQHVS-VHVLSSTT-ALMSWTSSQENYNSTIVSV-VSLTCQKQ 533

Qy        665 GNSSN-ATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMA 723
                    |     | :  :   :   :|:||: | | |: : |   |      |  |  || : |
Db        534 KESQRLEKQYCTQVNSSKPIIENLVPGAQYQVVIYLRKGPLI-----GPPS---DPVTFA 585

Qy        724 SF-----DCEVVPKEP-ALVLKWTCP-PGANAGFELEVSSGAWNNATHLESCSSENGTEY 776
                     | : |  | |:|| ||| |   |       : :|:    :| ||      :|| |  |
Db        586 IVPTGIKDLMLYPLGPTAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSEWTTYY 639

Qy        777 RTEVTYLNFSTS-----------YNISITTVSCGKMAAPTRNTCTTG----ITDPPPPDG 821
                  | :: : |             ||  :| |: |      | : | :       || |
Db        640 EIAAT-VSLTASVRIANLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAPV---- 691

Qy        822 SPNITSVSHNSVKVKFS------GFEASHGPIKAYAVI--------------LTTGEAGH 861
              :| |||| : :  :  |           : ||  :  : |:                : |
Db        692 APEITSVEYFNSLLYISWTYGDDTTDLSHSRMLHWMVVAEGKKKIKKSVTRNVMTAILSL 751

Qy        862 PSADVLKYTYDDFKKGASDTYYVTYLIRTE------------------------------ 890
               | |:    :       |:| :  |:: |
Db        752 PPGDIYNLSVTACTERGSNTSMLRLVKLEPAPPKSLFAVNKTQTSVTLLWVEEGVADFFE 811

Qy        891 ----EKGRSQSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKG 946
                    : ||          |: : |  :   |:         | |  :|     | |:
Db        812 VFCQQVGSSQKTK--LQEPVAVSSHVVTI----SSLLPATAYNCSVTSFSH--------- 856

Qy        947 LIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKK-----RK 1001
                 : |  :|   | |:   :| |: :|   :     |: : :    |   |||        |:
Db        857 ---DSPSVPTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQMARE 912

Qy       1002 DAKNNEVSFSQIKP-------------KKSKL---IRVENFEAYFKKQQADSNCGFAEEY 1045
                     |:|: ::                     || ||    :::::|:||  |       ||:   |: ::
Db        913 CGAGTFVNFASLERDGKLPYNWSKNGLKKRKLTNPVQLDDFDAYIKDMAKDSDYKFSLQF 972

Qy       1046 EDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKL-SVQTHSTDDYINANYMPGYH 1104
              |:|||:|:  | :||:|  || |||  |:||||  |||:| |:          ||||||:|||:
Db        973 EELKLIGLDIPHFAADLPLNRCKNRYTNILPYDFSRVRLVSMNEEEGADYINANYIPGYN 1032

Qy       1105 SKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQDYGD 1163
              | :::||||||||| |   |||:|| ::     |:|||:| |: | ||: |||  :::   |||
Db       1033 SPQEYIATQGPLPETRNDFWKMVLQQKSQIIVMLTQCNEKRRVKCDHYWPFTEEPIAYGD 1092

Qy       1164 ITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVP--DTTDLLINFRYL 1221
              ||| | ||     :|  | | :        |    :   |::|:|||||||   :   : :: | ::
Db       1093 ITVEMISEEEQDDWACRHFRIN--YADEMQDVMHFNYTAWPDHGVPTANAAESILQFVHM 1150

Qy       1222 VRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLM 1281
              ||       :|  : |:::||||||||||||||:|||:  |  :    ||: |:| ::| :|  |
Db       1151 VRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRDHEFVDILGLVSEMRSYRMSM 1208

Qy       1282 VQTEDQYVFLNQCV--LDIVRSQKDSKVDLIYQNTT 1315
              ||||:||:|::||| :  :  : |:   |:||:| :
Db       1209 VQTEEQYIFIHQCVQLMWMKKKQQFCISDVIYENVS 1244

RESULT 13
US-09-949-016-8055
; Sequence 8055, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8055
;   LENGTH: 1246
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8055

  Query Match             12.1%;  Score 854;  DB 4;  Length 1246;
  Best Local Similarity   28.0%;  Pred. No. 1.7e-46;
  Matches  279;  Conservative  158;  Mismatches  347;  Indels  212;  Gaps   41;

Qy        452 TPPVPVS-------DFRVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVEIT 499
              || :|         ||  :   ||      | |  |      :  |   :  |  || : |
Db        329 TPEIPSGNISSGWPDFNSSDYETTSQPYWWDSASAAPESEDEFVSVLPMEYENNSTLSET 388

Qy        500 TNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHV----------- 548
                  :       ||        : || |   :|                | ||:
Db        389 EKSTSGSFSFFPVQMILTWLPPKPPTAFDG-------------FHIHIEREENFTEYLM 434

Qy        549 VYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISP 608
              |      |   : | |     :|    :     | ||     | :|          :  |||
Db        435 VDEEAHEFVAELKEP---GKYKLSVTTFSSSGSCETRKSQSAKSL---------SFYISP 482

Qy        609 EVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCLLIEKA 664
                 :  | :  :      :| :|| : | ::|| | :||   :|::    :    |  :|
Db        483 SGE--WIEELT----EKPQHVS-VHVLSSTT-ALMSWTSSQENYNSTIVSV-VSLTCQKQ 533

Qy        665 GNSSN-ATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMA 723
                    |     | :  :   :   :|:||: | | |: : |   |      |  |  || : |
Db        534 KESQRLEKQYCTQVNSSKPIIENLVPGAQYQVVIYLRKGPLI-----GPPS---DPVTFA 585

Qy        724 SF-----DCEVVPKEP-ALVLKWTCP-PGANAGFELEVSSGAWNNATHLESCSSENGTEY 776
                     | : |  | |:|| ||| |   |       : :|:    :| ||      :|| |  |
Db        586 IVPTGIKDLMLYPLGPTAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSEWTTYY 639

Qy        777 RTEVTYLNFSTS-----------YNISITTVSCGKMAAPTRNTCTTG----ITDPPPPDG 821
                  | :: : |             ||  :| |: |      | : | :       || |
Db        640 EIAAT-VSLTASVRIANLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAPV---- 691
```

```
Qy        822 SPNITSVSHNSVKVKFS------GFEASHGPIKAYAVI-------------LTTGEAGH 861
Db        692 APEITSVEYFNSLLYISWTYGDDTTDLSHSRMLHWMVVAEGKKKIKKSVTRNVMTAILSL 751

Qy        862 PSADVLKYTYDDFKKGASDTYVTYLIRTE------------------------------ 890
Db        752 PPGDIYNLSVTACTERGSNTSMLRLVKLEPAPPKSLFAVNKTQTSVTLLWVEEGVADFFE 811

Qy        891 ----EKGRSQSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKG 946
Db        812 VFCQQVGSSQKTK--LQEPVAVSSHVVTI----SSLLPATAYNCSVTSFSH--------- 856

Qy        947 LIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKK-----RK 1001
Db        857 ---DSPSVPTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQMARE 912

Qy       1002 DAKNNEVSFSQIKP-------------KKSKL---IRVENFEAYFKKQQADSNCGFAEEY 1045
Db        913 CGAGTFVNFASLERDGKLPYNWSKNGLKKRKLTNPVQLDDFDAYIKDMAKDSDYKFSLQF 972

Qy       1046 EDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKL-SVQTHSTDDYINANYMPGYH 1104
Db        973 EELKLIGLDIPHFAADLPLNRCKNRYTNILPYDFSRVRLVSMNEEEGADYINANYIPGYN 1032

Qy       1105 SKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQDYGD 1163
Db       1033 SPQEYIATQGPLPETRNDFWKMVLQQKSQIIVMLTQCNEKRRVKCDHYWPFTEEPIAYGD 1092

Qy       1164 ITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVP--DTTDLLINFRYL 1221
Db       1093 ITVEMISEEEQDDWACRHFRIN--YADEMQDVMHFNYTAWPDHGVPTANAAESILQFVHM 1150

Qy       1222 VRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLM 1281
Db       1151 VRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRDHEFVDILGLVSEMRSYRMSM 1208

Qy       1282 VQTEDQYVFLNQCV--LDIVRSQKDSKVDLIYQNTT 1315
Db       1209 VQTEEQYIFIHQCVQLMWMKKKQQFCISDVIYENVS 1244
```

```
RESULT 14
US-08-201-697-2
; Sequence 2, Application US/08201697
; Patent No. 5705623
;  GENERAL INFORMATION:
;    APPLICANT:  Wiggins, Roger C.
;    APPLICANT:  Thomas, Peedikayil E.
;    TITLE OF INVENTION:  Mammalian Glomerular Epithelial Protein
;    TITLE OF INVENTION:  1
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/201,697
;      FILING DATE:  25-FEB-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryb A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UM 9783
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1187 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-201-697-2

  Query Match             12.0%;  Score 842;  DB 1;  Length 1187;
  Best Local Similarity   24.4%;  Pred. No. 9.3e-46;
  Matches  331;  Conservative  203;  Mismatches  436;  Indels  389;  Gaps   55;

Qy        134 NVILTWKSNDTAASEYKYVVK-------HKMENEKTITVVHQP---WCNITGLRPATSYV 183
Db         39 SIVVSLEASDVISPASVYVVKITGESKNYFFEFEEFNSTLPPPVIFKANYHGL----YYI 94

Qy        184 FSITPGIGNETWGDPRVIKVITEPIPVSDLRV------ALTGVRKAALSWSNGNGTASCR 237
Db         95 ITLVVVNGNVVTKPSRSITVLTKPLPVTSVSIYDYKPSPETGV---------------- 137

Qy        238 VLLESIGSHEELTQDSRLQVNISDLK--------------------PGVQYNINPYLLQ 276
Db        138 --LFEIHYPEKYNVFTRVNISYWEGKAFRTMLYKDFFKGKTVFNHWLPGICYSNITFQLV 195

Qy        277 SNKT--KGDPLGTEGGLDASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVG 334
Db        196 SEATFNKSTLVEYSG---VSHEPKQHRTAPYPP-----------------RNISVRIVN 234

Qy        335 LEPGTRYNATVYSQAANGTEGQ-----------PQAIEFRTNAIQVFDVTAVNISATSL 382
Db        235 LN------------KNNWEEQSGSFPEESFMRSPETIE----KDRIF------------ 265

Qy        383 TLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAVIPGLRSSTFYNITVCPVLGDIE 442
Db        266 ------------------HFTEETPEPSGNIS-------------------------S 280

Qy        443 GTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVE 497
Db        281 GWP------------DFNSSDYETTSQPYWWDSASATPESEDEFVSVLPMEYENNNTLS 327

Qy        498 ITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHVVYVTTTEMW 557
Db        328 EAEKPTPAPFSFFPVQMILSWLPPKPPTAFDG-------------FHIH---IEREENF 370

Qy        558 LDWKSPD-GASEYVYHLVIESKH----------GSNHTSTYDKAITLQGLIPGT---LYN 603
Db        371 TEYSTVDEEAHEFVAELKEPGKYKLSVTTFSASGSCETRESQSAKSLSFYISPTGEWIEE 430

Qy        604 ITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCL 659
Db        431 LTEKPQ--HV-----------------SVHVLSSTT-ALMSWTSSQENYNSTIVSV-VSL 469

Qy        660 LIEKAGNSSN-ATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTD 718
Db        470 TCQKQKESQRLEKQYCTQVNSSKRIIENLVPGAQYQVVMYLRKGPLI-----GPPS---D 521

Qy        719 PASMASF-----DCEVVPKEP-ALVLKWTCP-PGANAGPELEVSSGAWNNATHLESCSSE 771
Db        522 PVTFAIVPTGIKDLMLYPLGPTAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSE 575

Qy        772 NGTEYRTEVTYLNFSTS-----------YNISITTVSCGKMAAPTRNTCTTG----ITDP 816
Db        576 WTTYYEIAAT-VSLTASVRIANLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAP 631

Qy        817 PPPDGSPNITSVSHNSVKVKFS------GFEASHGPIKAYAVI-------------LTT 856
Db        632 V----APEITSVEYFNSLLYISWTYGDDTTDLSHSRMLHWMVVAEGKKKIKKSVTRNVMT 687

Qy        857 GEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTE--------------------EKGR 894
Db        688 AILSLPPGDIYNLSVTACTERGSNTSMLRLVKLEPAPPKSLFAVNKTQTSVTLLWVEEGV 747
```

```
Qy        895 S-----------QSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQ 943
Db        748 ADFFEVFCQQVGSGLETKLQEPVAVSSHVVTI----SSLLPATAYNCSVTSFSH------ 797

Qy        944 NKGLIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKK---- 999
Db        798 ------DSPSVPTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQM 850

Qy       1000 -RKDAKNNEVSFSQIKP-------------KKSKL---IRVENFEAYFKKQQADSNCGFA 1042
Db        851 ARECGAGTFVNFASLERDGKLPYNWSKNGLKKRKLTNPVQLDDFDAYIKDMAKDSDYKFS 910

Qy       1043 EEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVK-LSVQTHSTDDYINANYMP 1101
Db        911 LQFEELKLIGLDIPHFAADLPLNRCKNRYTNILPYDFSRVRLLSMNEEEGADYINANYIP 970

Qy       1102 GYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQD 1160
Db        971 GYNSPQEYIATQGPLPETRNDFWKMVLQQKSQMIVMLTQCNEKRRVKCDHYWPFTEEPIA 1030

Qy       1161 YGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVP--DTTDLLINF 1218
Db        1031 YGDITVEMISEEEQDDWAHRHFRIN--YADEMQDVMHFNYTAWPDHGVPTANAAESILQF 1088

Qy       1219 RYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHR 1278
Db        1089 VHMVRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRDHEFVDILGLVSEMRSYR 1146

Qy       1279 PLMVQTEDQYVFLNQCV--LDIVRSQKDSKVDLIYQNTT 1315
Db        1147 MSMVQTEEQYIFIHQCVQLMWMKKKQQFCISDVIYENVS 1185
```

RESULT 15
US-09-949-016-8828
; Sequence 8828, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8828
; LENGTH: 1274
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8828

Query Match 12.0%; Score 841; DB 4; Length 1274;
Best Local Similarity 27.2%; Pred. No. 1.2e-45;
Matches 279; Conservative 159; Mismatches 346; Indels 240; Gaps 43;

```
Qy        452 TPPVPVS-------DFRVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVEIT 499
Db        329 TPEIPSGNISSGWPDFNSSDYETTSQPYWWDSASAAPESEDEFVSVLPMEYENNSTLSET 388

Qy        500 TNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHV----------- 548
Db        389 EKSTSGSFSFFPVQMILTWLPPKPPTAFDG-------------FHIHIEREENFTEYLM 434

Qy        549 VYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISP 608
Db        435 VDEEAHEFVAELKEP---GKYKLSVTTFSSSGSCETRKSQSAKSL---------SFYISP 482

Qy        609 EVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCLLIEKA 664
Db        483 SGE--WIEELT----EKPQHVS-VHVLSSTT-ALMSWTSSQENYNSTIVSV-VSLTCQKQ 533

Qy        665 GNSSN-ATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMA 723
Db        534 KESQRLEKQYCTQVNSSKPIIENLVPGAQYQVVIYLRKGPLI-----GPPS---DPVTFA 585

Qy        724 SF-----DCEVVPKEP-ALVLKWTCP-PGANAGFELEVSSGAWNNATHLESCSSENGTEY 776
Db        586 IVPTGIKDLMLYPLGPTAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSEWTTYY 639

Qy        777 RTEVTYLNFSTS-----------YNISITTVSCGKMAAPTRNTCTTG----ITDPPPPDG 821
Db        640 EIAAT-VSLTASVRIANLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAPV---- 691

Qy        822 SPNITSVSHNSVKVKFS------GFEASHGPIKAYAVI--------------LTTGEAGH 861
Db        692 APEITSVEYFNSLLYISWTYGDDTTDLSHSRMLHWMVVAEGKKKIKKSVTRNVMTAILSL 751

Qy        862 PSADVLKYTYDDFKKGASDTYVTYLIRTE------------------------------ 890
Db        752 PPGDIYNLSVTACTERGSNTSMLRLVKLEPAPPKSLFAVNKTQTSVTLLWVEEGVADFFE 811

Qy        891 ----EKGRSQSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKG 946
Db        812 VFCQQVGSSQKTK--LQEPVAVSSHVVTI----SSLLPATAYNCSVTSFSH--------- 856

Qy        947 LIDGAESYVSFSRYSDAVSLPQDPGVICGAVFG----CIFGALVIVTV------------ 990
Db        857 ---DSPSVPTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQMARE 912

Qy        991 ---GGFIFWRKKRKDAK------NNEVSFSQIKP--------------------KKSKL 1020
Db        913 CGAGTFVNFASLERDGKLPYNWRRSIFAFLTLLPSCLWTDYLLAFYINPWSKNGLKKRKL 972

Qy       1021 ---IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPY 1077
Db        973 TNPVQLDDFDAYIKDMAKDSDYKFSLQFEELKLIGLDIPHFAADLPLNRCKNRYTNILPY 1032

Qy       1078 DISRVKL-SVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAII 1136
Db        1033 DFSRVRLVSMNEEEGADYINANYIPGYNSPQEYIATQGPLPETRNDFWKMVLQQKSQIIV 1092

Qy       1137 MLTKCVEQGRTKCEEYWP-SKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPL 1195
Db        1093 MLTQCNEKRRVKCDHYWPFTEEPIAYGDITVEMISEEEQDDWACRHFRIN--YADEMQDV 1150

Qy       1196 RQFHFTSWPDHGVP--DTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAID 1253
Db        1151 MHFNYTAWPDHGVPTANAAESILQFVHMVRQQATKS--KGPMIIHCSAGVGRTGTFIALD 1208

Qy       1254 RLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCV--LDIVRSQKDSKVDLIY 1311
Db        1209 RLLQHIRDHEFVDILGLVSEMRSYRMSMVQTEEQYIFIHQCVQLMWMKKKQQFCISDVIY 1268

Qy       1312 QNTT 1315
Db        1269 ENVS 1272
```

Search completed: May 10, 2005, 19:06:31
Job time : 60 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 10, 2005, 18:57:57 ; Search time 168 Seconds
(without alignments)
2654.891 Million cell updates/sec

Title: US-09-516-728A-4
Perfect score: 7030
Sequence: 1 MKPAAREARLPPRSPGLRWA.....TIYENLAPVTTFGKTNGYIA 1337

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1428581 seqs, 333598853 residues

Total number of hits satisfying chosen parameters: 1428581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7030	100.0	1337	15	US-10-366-547-42	Sequence 42, Appl
2	7030	100.0	1337	16	US-10-723-606-2	Sequence 2, Appli
3	6950	98.9	1337	15	US-10-366-547-44	Sequence 44, Appl
4	6824	97.1	1337	14	US-10-390-501-2	Sequence 2, Appli
5	4508.5	64.1	1238	15	US-10-366-547-47	Sequence 47, Appl
6	4499.5	64.0	1216	15	US-10-366-547-49	Sequence 49, Appl
7	1811	25.8	341	16	US-10-723-606-3	Sequence 3, Appli
8	1242	17.7	1997	15	US-10-634-027-2	Sequence 2, Appli
9	1241	17.7	1997	10	US-09-909-567B-54	Sequence 54, Appl
10	1241	17.7	1997	16	US-10-408-765A-2135	Sequence 2135, Ap
11	1241	17.7	1997	16	US-10-497-692-4	Sequence 4, Appli
12	1230	17.5	1450	16	US-10-497-692-14	Sequence 14, Appl
13	1130	16.1	1118	14	US-10-245-539-2	Sequence 2, Appli
14	1130	16.1	1118	14	US-10-245-539-8	Sequence 8, Appli
15	1127	16.0	1093	14	US-10-245-539-4	Sequence 4, Appli
16	969	13.8	1767	15	US-10-087-684-40	Sequence 40, Appl
17	969	13.8	1767	15	US-10-218-779-40	Sequence 40, Appl
18	954	13.6	1767	15	US-10-087-684-41	Sequence 41, Appl
19	954	13.6	1767	15	US-10-218-779-41	Sequence 41, Appl
20	924.5	13.2	1447	16	US-10-497-692-13	Sequence 13, Appl
21	889.5	12.7	336	15	US-10-634-027-4	Sequence 4, Appli
22	879	12.5	1912	17	US-10-772-636-64	Sequence 64, Appl
23	878.5	12.5	312	15	US-10-634-027-6	Sequence 6, Appli
24	878.5	12.5	319	15	US-10-634-027-7	Sequence 7, Appli
25	864	12.3	2301	10	US-09-822-871-4	Sequence 4, Appli
26	864	12.3	2301	15	US-10-673-885-4	Sequence 4, Appli
27	864	12.3	2302	15	US-10-087-684-37	Sequence 37, Appl
28	864	12.3	2302	15	US-10-218-779-37	Sequence 37, Appl
29	854	12.1	1188	15	US-10-331-496A-79	Sequence 79, Appl
30	854	12.1	1188	16	US-10-789-241-34	Sequence 34, Appl
31	831.5	11.8	2281	15	US-10-087-684-6	Sequence 6, Appli
32	831.5	11.8	2281	15	US-10-218-779-6	Sequence 6, Appli
33	827	11.8	2299	15	US-10-466-759-2	Sequence 2, Appli
34	826	11.7	2291	10	US-09-822-871-2	Sequence 2, Appli
35	826	11.7	2291	15	US-10-673-885-2	Sequence 2, Appli
36	826	11.7	2300	15	US-10-087-684-10	Sequence 10, Appl
37	826	11.7	2300	15	US-10-218-779-10	Sequence 10, Appl
38	817	11.6	1367	15	US-10-369-493-5508	Sequence 5508, Ap
39	817	11.6	1367	15	US-10-369-493-5509	Sequence 5509, Ap
40	799	11.4	1948	9	US-09-808-602-55	Sequence 55, Appl
41	799	11.4	1948	10	US-09-800-198-45	Sequence 45, Appl
42	792.5	11.3	1897	17	US-10-482-029-52	Sequence 52, Appl
43	788.5	11.2	1907	15	US-10-291-265-250	Sequence 250, App
44	782	11.1	310	9	US-09-788-626-15	Sequence 15, Appl
45	778.5	11.1	1502	9	US-09-808-602-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-10-366-547-42
; Sequence 42, Application US/10366547
; Publication No. US20030215899A1
; GENERAL INFORMATION:
; APPLICANT: Meng, Tzu-Ching
; APPLICANT: Tonks, Nicholas K.
; APPLICANT: Cool, Deborah E.
; TITLE OF INVENTION: REVERSIBLE OXIDATION OF PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.439
; CURRENT APPLICATION NUMBER: US/10/366,547
; CURRENT FILING DATE: 2003-02-12
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-366-547-42

Query Match 100.0%; Score 7030; DB 15; Length 1337;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
```

```
Qy       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
Db       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
Db       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy       301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
Db       301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy       361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
Db       361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy       421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
Db       421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy       481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy       541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
Db       541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy       601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
Db       601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy       661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
Db       661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720

Qy       721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
Db       721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy       781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
Db       781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy       841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
Db       841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy       901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
Db       901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy       961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
Db       961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy      1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
Db      1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy      1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
Db      1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy      1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
Db      1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy      1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
Db      1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy      1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
Db      1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy      1321 ENLAPVTTFGKTNGYIA 1337
Db      1321 ENLAPVTTFGKTNGYIA 1337
```

RESULT 2
US-10-723-606-2
; Sequence 2, Application US/10723606
; Publication No. US20040161821A1
; GENERAL INFORMATION:
; APPLICANT: Palka-Hamblin, Helena L.
; APPLICANT: Tonks, Nicholas K.
; TITLE OF INVENTION: DEP-1 RECEPTOR PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE INTERACTING PROTEINS
; TITLE OF INVENTION: AND RELATED METHODS
; FILE REFERENCE: 200125.447
; CURRENT APPLICATION NUMBER: US/10/723,606
; CURRENT FILING DATE: 2003-11-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-606-2

Query Match 100.0%; Score 7030; DB 16; Length 1337;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
Db         1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy        61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
Db        61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy       121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
Db       121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
Db       181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
Db       241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy       301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
Db       301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy       361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
Db       361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy       421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
Db       421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy       481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy       541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
Db       541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
```

```
Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337
```

```
RESULT 3
US-10-366-547-44
; Sequence 44, Application US/10366547
; Publication No. US20030215899A1
; GENERAL INFORMATION:
;  APPLICANT: Meng, Tzu-Ching
;  APPLICANT:  Tonks, Nicholas K.
;  APPLICANT:  Cool, Deborah E.
;  TITLE OF INVENTION: REVERSIBLE OXIDATION OF PROTEIN TYROSINE
;  TITLE OF INVENTION:  PHOSPHATASES
;  FILE REFERENCE: 200125.439
;  CURRENT APPLICATION NUMBER: US/10/366,547
;  CURRENT FILING DATE:  2003-02-12
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
;   LENGTH: 1337
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

US-10-366-547-44

Query Match 98.9%; Score 6950; DB 15; Length 1337;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1325; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISGLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              |||||||||||||||||   :    |  ||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYLQWEAGTSGLLPACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
```

```
Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
             |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337
```

RESULT 4
US-10-390-501-2
; Sequence 2, Application US/10390501
; Publication No. US20030148491A1
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.
; Ostman, Arne
; TITLE OF INVENTION: DENSITY ENHANCED PROTEIN TYROSINE
; PHOSPHATASES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group PLLC
; STREET: Suite 6300, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 980104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/390,501
; FILING DATE: 13-Mar-2003
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 200125.402C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-390-501-2

Query Match 97.1%; Score 6824; DB 14; Length 1337;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1302; Conservative 6; Mismatches 29; Indels 0; Gaps 0;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVAHGCEEGCSLSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLAQKVAWMPAIQREAG 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 QGAPPPLCMMSPFVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVELFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337


RESULT 5
US-10-366-547-47
; Sequence 47, Application US/10366547
; Publication No. US20030215899A1
; GENERAL INFORMATION:
;  APPLICANT: Meng, Tzu-Ching
;  APPLICANT:  Tonks, Nicholas K.
;  APPLICANT:  Cool, Deborah E.
;  TITLE OF INVENTION: REVERSIBLE OXIDATION OF PROTEIN TYROSINE
;  TITLE OF INVENTION:  PHOSPHATASES
;  FILE REFERENCE: 200125.439
;  CURRENT APPLICATION NUMBER: US/10/366,547
;  CURRENT FILING DATE:  2003-02-12
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
;   LENGTH: 1238
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-366-547-47

  Query Match             64.1%;  Score 4508.5;  DB 15;  Length 1238;
  Best Local Similarity   67.2%;  Pred. No. 2.1e-268;
  Matches  901;  Conservative  108;  Mismatches  226;  Indels  105;  Gaps   12;

Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||| | |||||||||||  |||||| ||:||||  |:||
Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLRQGQVLCAGAAPNPI----------------- 43

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Db         44 ------------------------------------------------------------ 43

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
                |||:|| :|||:|:|||| ||: ||| :  :::|||:  |  | :|  |||||| | |
Db         44 --FDIEAV-VSPTSVLLTWKHNDSGASECR--IENKMESNLTFPVKNQTSCNITGLSPGT 98

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              || |||     |||        : ||| ||||| |   || :| |:||| ||||| |:|:
Db         99 SYTFSIISVTTNETLNK----TITTEPWPVSDLHVTSVGVTQARLTWSNANGTASYRMLI 154

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLL-QSNKTKGDPLGTEGGLDASNTERS 299
              | : :|          |||| ||||     | :   :||:|: |    |   ||: |:|
Db        155 EELTTHS--------SVNISGLKPGTN---NTFAFPESNETQADFAVAEEVPDANGTKRI 203

Qy        300 RAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQA 359
                :  : :|   |||  ||| |||    ||:||  |:| |:||||:||||||||||:
Db        204 PV-TNLSQLHKNSLVS-VDPPSGQDPSLTEILLTDLKPDTQYNATIYSQAANGTEGQPRN 261

Qy        360 IEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419
              |:||: || || |:||||:|:|| || : : | ::  ||||||| | | |  |::  |
Db        262 KVFKTNSTQVSDVRAMNISASSMTLTWKSNYDGSRTSIVYKIHVAGGTHSVNQTVNKTEA 321

Qy        420 VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA 479
                :| || ||| ||||| | ||  |||||||||:| |  ||||||| |||  |||||  |:|:
Db        322 IILGLSSSTLYNITVHPFLGQTEGTPGFLQVYTSPDQVSDFRVTNVSTRAIGLAWRSNDS 381

Qy        480 ESFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTV 539
              :||:: | |:|    |    | ||| ::  | ||| | |||:|:||:|||| | ||   |
Db        382 KSFEIFIKQDGGEKHRNASTGNQSYMVEDLKPGTSYHFEIIPRGPDGTEGLSSTVNGSTD 441

Qy        540 PSAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGS-NHTSTYDKAITLQGLIP 598
              |||| || || ::|||| |:|:: | || | ||||:||| ||    |::  | ||:  | |
Db        442 PSAVTDIRVVNISTTEMQLEWQNTDDASGYTYHLVLESKSGSIIRTNSSQKWITVGSLTP 501

Qy        599 GTLYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYC 658
              |||||:|| |||| : |  ||  ||||||:||:|:|:| || | : |:|  | || :|:|
Db        502 GTLYNVTIFPEVDQIQGISNSITQYTRPSSVSHIEVNTTTTTAAIRWKNEDAASASYAYS 561

Qy        659 LLIEKAGNSSNATQVVT-DIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSPCT 717
              :|| | |: || |   | |  |   : ||||| ||||:|  |||||  || ||   |||
Db        562 VLILKTGDGSNVTSNFTKDPSI---LIPELIPGVSYTVKILTQVGDGTTSLVPGWNLFCT 618

Qy        718 DPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYR 777
              :|  : || |||||||||||||| || |   |||| | | :|:| | ||:|:|:: || |
Db        619 EPEPVTSFHCEVVPKEPALVLKWACPFGMYTGFELGVRSDSWDNMTRLENCTSDDDTECR 678

Qy        778 TEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKF 837
              ||| |||||||||||| |:|||||| | :| |||||||| ||||||||||||||||||||
Db        679 TEVAYLNFSTSYNISIATLSCGKMALPAQNICTTGITDPPTPDGSPNITSVSHNSVKVKF 738

Qy        838 SGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQS 897
              |||||||||||||||||||||| |||||||||:|||:|||||||||| ||||:||
Db        739 SGFEASHGPIKAYAVILTTGEAAQPSADVLKYTYEDFKRGASDTYVTYLIRIEEKGQSQG 798

Qy        898 LSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSF 957
              ||||| |||||||:|||||||||:|||||||||||||||||:: || |||:| ||||||
Db        799 LSEVLNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYVSF 858

Qy        958 SRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKK 1017
              | ||:|| ||||||||||||||||||| |  |||||||||| ||||||||||||||||||
Db        859 SPYSEAVFLPQDPGVICGAVFGCIFGALAITAVGGFIFWRKKRTDAKNNEVSFSQIKPKK 918

Qy       1018 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPY 1077
              |||||||||||||||||||||||||||||||||:||| ||| ||:|||||||||||||||
Db        919 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYTAEIAENRGKNRYNNVLPY 978

Qy       1078 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIM 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        979 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIVM 1038

Qy       1138 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQ 1197
              |||||||||||||||||||||||||||||||||:|||||||| |||:| ||||||||
Db       1039 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEVVLPEWTIRDFVVKNMQNSESHPLRQ 1098

Qy       1198 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIY 1257
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRLIY 1158

Qy       1258 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAM 1317
              |||||||||||||||||||||||||||||||||||||||:|:|||||||||||||||||
Db       1159 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTTAM 1218

Qy       1318 TIYENLAPVTTFGKTNGYIA 1337
              |||||| ||: |||||||||
Db       1219 TIYENLEPVSMFGKTNGYIA 1238


RESULT 6
US-10-366-547-49
; Sequence 49, Application US/10366547
; Publication No. US20030215899A1
; GENERAL INFORMATION:
;  APPLICANT: Meng, Tzu-Ching
;  APPLICANT:  Tonks, Nicholas K.
;  APPLICANT:  Cool, Deborah E.
;  TITLE OF INVENTION: REVERSIBLE OXIDATION OF PROTEIN TYROSINE
```

```
;  TITLE OF INVENTION:  PHOSPHATASES
;  FILE REFERENCE: 200125.439
;  CURRENT APPLICATION NUMBER: US/10/366,547
;  CURRENT FILING DATE:  2003-02-12
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
;   LENGTH: 1216
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-10-366-547-49

  Query Match             64.0%;  Score 4499.5;  DB 15;  Length 1216;
  Best Local Similarity   67.1%;  Pred. No. 7.3e-268;
  Matches  900;  Conservative  106;  Mismatches  205;  Indels  131;  Gaps   10;

Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLRQGQVVCTG---------------------- 37

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
Db         38 ------------------------------------------------------AAPS 41

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
Db         42 PVFDVEAVT-SPTSVVLTWKHNDSATSEYKI----NEGNTLRYTVKNQTSFNITGLSPAT 96

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
Db         97 SYKFSITLGTVNETSGKPTYKNITTEPWPVSDLQVAYIGVTQALLAWSNANGTASYRMQI 156

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
Db        157 ------VELTTNS--SGGISDLKPGTHKSL--AVQGSNETQHDLWVTE------------ 194

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
Db        195 -------------GVSDPPSARDPSLTEVLLTELKPDTQYKVTIYSQAADGTEGQPGNK 240

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
Db        241 VFKTNPIQVSDIRAVNISDSNMTLTWKSNNNESHASFTYKIYVAGGSDSINETVNETQAV 300

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
Db        301 IRGLSSSTLYNITVLPFLGQTAGIPGFLQVYTSPRPVSDFRVTNVSLREIGLAWRSNDSE 360

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db        361 SFEIFITQEGSEKRWNASTGDLSYIVVNLKPGTSYQFEIFPRGPNGTEGPSQTVVGRTDC 420

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
Db        421 SAVTDIRVVSVSTTEIQLEWQNTDSASGYTYHLVLESENGSIKTNSSQKWITVGGLTPGT 480

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNPDDASPTYSYCLL 660
Db        481 LYNVTIFPEVDQMEGNSSSITQYTRPSNVSYIEVNTNTTVGAIQWKNLDAASASYSYSVL 540

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIF-----AQVGDGIKSLEPGRKSF 715
Db        541 ILKAGDGSNVTSRVRD--IPSVTIPGLIPGVSSEVKIFTKIRNTEVGNEV----PGQKLF 594

Qy        716 CTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTE 775
Db        595 CMEPAQVDSLHCEVVPKEPALVLKWACPPGMNSGFELGVRSDAWDNMTHLENCTLDNDTE 654

Qy        776 YRTEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKV 835
Db        655 CRTEVTYLNFSTSYNISIATLSCGKMALPTQSTCTTGITDPPPPDGSPNITSVSHNSVKV 714

Qy        836 KFSGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRS 895
Db        715 KFSGFEASHGPIKAYAVLLTTGEAGQPSTDVLKYTYEDFKKGASDTYVTYLIRIEEKGQS 774

Qy        896 QSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYV 955
Db        775 QGLSEALNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYV 834

Qy        956 SFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKP 1015
Db        835 SFSPYSEAVSLPQDPGVICGAVFGCIFGALAIVAVGGFIFWRKKRKDAKNNEVSFSQIKP 894

Qy       1016 KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVL 1075
Db        895 KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYAAEIAENRGKNRYNNVL 954

Qy       1076 PYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1135
Db        955 PYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1014

Qy       1136 IMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPL 1195
Db       1015 VMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEVVLPEWTIRDFVVKNMQSSESHPL 1074

Qy       1196 RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRL 1255
Db       1075 RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRL 1134

Qy       1256 IYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTT 1315
Db       1135 IYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTT 1194

Qy       1316 AMTIYENLAPVTTFGKTNGYIA 1337
Db       1195 AMTIYENLERVSMFGKANGYIA 1216


RESULT 7
US-10-723-606-3
; Sequence 3, Application US/10723606
; Publication No. US20040161821A1
; GENERAL INFORMATION:
;  APPLICANT: Palka-Hamblin, Helena L.
;  APPLICANT:  Tonks, Nicholas K.
;  TITLE OF INVENTION: DEP-1 RECEPTOR PROTEIN TYROSINE
;  TITLE OF INVENTION:  PHOSPHATASE INTERACTING PROTEINS
;  TITLE OF INVENTION:  AND RELATED METHODS
;  FILE REFERENCE: 200125.447
;  CURRENT APPLICATION NUMBER: US/10/723,606
;  CURRENT FILING DATE:  2003-11-26
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 341
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-723-606-3

  Query Match             25.8%;  Score 1811;  DB 16;  Length 341;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-103;
  Matches  341;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        997 RKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQP 1056
Db          1 RKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQP 60

Qy       1057 KYAAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPL 1116
Db         61 KYAAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPL 120

Qy       1117 PNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPE 1176
Db        121 PNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPE 180
```

```
Qy       1177 WTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPIL 1236
Db        181 WTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPIL 240

Qy       1237 VHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVL 1296
Db        241 VHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVL 300

Qy       1297 DIVRSQKDSKVDLIYQNTTAMTIYENLAPVTTFGKTNGYIA 1337
Db        301 DIVRSQKDSKVDLIYQNTTAMTIYENLAPVTTFGKTNGYIA 341


RESULT 8
US-10-634-027-2
; Sequence 2, Application US/10634027
; Publication No. US20040077065A1
; GENERAL INFORMATION:
;  APPLICANT: Procter & Gamble Company
;  APPLICANT:  Evdokimov, Artem G
;  APPLICANT:  Pokross, Matthew E
;  TITLE OF INVENTION: Three Dimensional Coordinates of HPTPbeta
;  FILE REFERENCE: 9045M2
;  CURRENT APPLICATION NUMBER: US/10/634,027
;  CURRENT FILING DATE:  2003-08-04
;  PRIOR APPLICATION NUMBER: US 60/413,547
;  PRIOR FILING DATE: 2002-09-25
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 1997
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-634-027-2

  Query Match             17.7%;  Score 1242;  DB 15;  Length 1997;
  Best Local Similarity   29.6%;  Pred. No. 4.5e-67;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
Db        759 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL-----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
Db        845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG-----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
Db       1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKRDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325
Db       1970 RS----EQENPLFPIYENVNP 1986


RESULT 9
US-09-909-567B-54
; Sequence 54, Application US/09909567B
; Publication No. US20030022257A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto A.
;  APPLICANT:  Nair, Manoj
;  APPLICANT:  Chen, Seiyu
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes
;  FILE REFERENCE: DEX-0214
;  CURRENT APPLICATION NUMBER: US/09/909,567B
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: 60/219,834
;  PRIOR FILING DATE: 2000-07-21
```

```
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
;   LENGTH: 1997
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-909-567B-54

  Query Match             17.7%;  Score 1241;  DB 10;  Length 1997;
  Best Local Similarity   29.6%;  Pred. No. 5.1e-67;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy       123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
Db       759 FDHYEVTIKNKNNFIQTKSIP----------KSENE----------CVFVQLVPGRLY 796

Qy       183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
Db       797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy       239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
Db       845 ---DVDQYEIQLLFNDMKVFPPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy       281 KG----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
Db       902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy       314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
Db       958 -------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy       374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
Db      1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy       425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
Db      1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy       485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db      1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy       541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
Db      1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy       597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
Db      1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy       648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
Db      1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy       705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
Db      1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy       764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
Db      1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy       819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
Db      1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy       863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
Db      1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKCDPT 1562

Qy       917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967

Db      1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy       968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
Db      1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy      1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
Db      1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy      1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
Db      1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy      1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
Db      1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy      1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
Db      1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy      1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
Db      1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy      1305 SKVDLIYQNTTAMTIYENLAP 1325
Db      1970 RS----EQENPLFPIYENVNP 1986


RESULT 10
US-10-408-765A-2135
; Sequence 2135, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
;  APPLICANT: Ghosh, Soumitra S.
;  APPLICANT:  Fahy, Eoin D.
;  APPLICANT:  Zhang, Bing
;  APPLICANT:  Gibson, Bradford W.
;  APPLICANT:  Taylor, Steven W.
;  APPLICANT:  Glenn, Gary M.
;  APPLICANT:  Warnock, Dale E.
;  TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;  TITLE OF INVENTION:  IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;  FILE REFERENCE: 660088.465
;  CURRENT APPLICATION NUMBER: US/10/408,765A
;  CURRENT FILING DATE:  2003-04-04
;  NUMBER OF SEQ ID NOS: 3077
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2135
;   LENGTH: 1997
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-408-765A-2135

  Query Match             17.7%;  Score 1241;  DB 16;  Length 1997;
  Best Local Similarity   29.6%;  Pred. No. 5.1e-67;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy       123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
Db       759 FDHYEVTIKNKNNFIQTKSIP----------KSENE----------CVFVQLVPGRLY 796

Qy       183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
Db       797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy       239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
Db       845 ---DVDQYEIQLLFNDMKVFPPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901
```

```
Qy        281 KG------------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
Db       1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKCDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
```

```
Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325
Db       1970 RS----EQENPLFPIYENVNP 1986
```

RESULT 11
US-10-497-692-4
; Sequence 4, Application US/10497692
; Publication No. US20050004056A1
; GENERAL INFORMATION:
; APPLICANT: Meise, Martin
; APPLICANT: Eulenberg, Karsten
; APPLICANT: Fritsch, Rudiger
; APPLICANT: Hader, Thomas
; APPLICANT: Bronner, Gunter
; APPLICANT: Steuernagel, Arnd
; TITLE OF INVENTION: PTP10D, Tec protein tyrosine kinase and EDTP homologous proteins
; TITLE OF INVENTION: involved in the regulation of energy homeostasis
; FILE REFERENCE: 2923-632
; CURRENT APPLICATION NUMBER: US/10/497,692
; CURRENT FILING DATE: 2004-06-04
; PRIOR APPLICATION NUMBER: PCT/EP02/13744
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: EP 01 000 010.5
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: EP 01 129 138.2
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: EP 01 128 844.6
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1997
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-497-692-4

Query Match 17.7%; Score 1241; DB 16; Length 1997;
Best Local Similarity 29.6%; Pred. No. 5.1e-67;
Matches 397; Conservative 190; Mismatches 503; Indels 251; Gaps 53;

```
Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
Db        759 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
Db        845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG------------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
```

```
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596

Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647

Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704

Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763

Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818

Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862

Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916

Db       1509 LPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLGAEMESLGGKCDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967

Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015

Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067

Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126

Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185

Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244

Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304

Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325

Db       1970 RS----EQENPLFPIYENVNP 1986
```

RESULT 12
US-10-497-692-14
; Sequence 14, Application US/10497692
; Publication No. US20050004056A1
; GENERAL INFORMATION:
; APPLICANT: Meise, Martin
; APPLICANT: Eulenberg, Karsten
; APPLICANT: Fritsch, Rudiger
; APPLICANT: Hader, Thomas
; APPLICANT: Bronner, Gunter
; APPLICANT: Steuernagel, Arnd
; TITLE OF INVENTION: PTP10D, Tec protein tyrosine kinase and EDTP homologous proteins
; TITLE OF INVENTION: involved in the regulation of energy homeostasis
; FILE REFERENCE: 2923-632
; CURRENT APPLICATION NUMBER: US/10/497,692
; CURRENT FILING DATE: 2004-06-04
; PRIOR APPLICATION NUMBER: PCT/EP02/13744
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: EP 01 000 010.5
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: EP 01 129 138.2
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: EP 01 128 844.6
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1450
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-497-692-14

Query Match 17.5%; Score 1230; DB 16; Length 1450;
Best Local Similarity 29.3%; Pred. No. 1.5e-66;
Matches 390; Conservative 197; Mismatches 495; Indels 250; Gaps 53;

```
Qy        125 IKAVSISPT----NVILTWKSNDTAASEYKYVVKH-----------KMENEKTITVVHQP 169

Db        216 VQGVSVSNSARSDYLRVSWVHATGDFDHYEVTIKNKNNFIQTKSIPKSENE--------- 266

Qy        170 WCNITGLRPATSYVFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAAL 225

Db        267 -CVFVQLVPGRLYSVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHV 319

Qy        226 SWSNGNGTASCRVLLESIGSHE--ELTQDSRL-----------QVNISDLKPGVQY---- 268

Db        320 TWSGANG---------DVDQYEIQLLFNDMKVFPPPFHLVNTATEYRFTSLTPGRQYKILV 370

Qy        269 -NINPYLLQSNKTKG-----------DPLG---------TEGGLDASN------TERSR 300

Db        371 LTISGDVQQSAFIEGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQK 430

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Db        431 VDSQTIPKH----------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINV 471

Qy        361 EFRTNAIQVFDVTAVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419

Db        472 VGRTVPASVQGVIADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPAT 528

Qy        420 --------VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIG 471

Db        529 TKQHKFEDLTPGKK----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLS 584

Qy        472 LAWSSHDAESFQMHI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGT 527

Db        585 FRWTASEGELSWYNIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGE 642

Qy        528 EGASRTVCNRTVPSAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTST 586

Db        643 LSNESFIFGRTVPASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENW 701

Qy        587 YDKAIT---LQGLIPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNT 638

Db        702 KDKDLTEWRFQGLVPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTS 758

Qy        639 TAATL----SWQNFDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGS 691

Db        759 LAITWKGPPDWTDYNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGR 808

Qy        692 SYTVEIFAQVGDGIKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGF 750
```

```
Db        809 SYQFNVKTVSGDSWKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGY 867

Qy        751 ELEVSSGAWNNATHLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPT 805

Db        868 SIE--------CRKMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVV 919

Qy        806 RNTCTTGITDPPPPDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG 857

Db        920 EDSTITMIDRPPPP--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREA 977

Qy        858 --------EAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVG 909

Db        978 DGSDELKPEQQHPLPSYLEYRHNASIRVYQTNY--FASKCAENPNSNSKS----FNIKLG 1031

Qy        910 NESTTLG---------YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR- 959

Db       1032 AEMESLGGKCDPTQQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKP 1080

Qy        960 -YSDA-VSLP----QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNE 1007

Db       1081 LYSDTFFSLPITTESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSAR 1138

Qy       1008 VSFSQIKP-------------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGIS 1054

Db       1139 LSIRRDRPLSVHLNLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRN 1198

Qy       1055 QPKYAAELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQ 1113

Db       1199 QSCDIALLPENRGKNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQ 1258

Qy       1114 GPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEI 1172

Db       1259 GPLPGTKDDFWKMVWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSES 1318

Qy       1173 VLPEWTIRDFTVKNIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPP 1231

Db       1319 VLPEWTIREFKICGEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPG 1378

Qy       1232 ESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFL 1291

Db       1379 AGPTVVHCSAGVGRTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYL 1438

Qy       1292 NQCVLDIVRSQK 1303

Db       1439 HQCVRDVLRARK 1450


RESULT 13
US-10-245-539-2
; Sequence 2, Application US/10245539
; Publication No. US20030077638A1
; GENERAL INFORMATION:
;  APPLICANT: Logan, Thomas Joseph
;  TITLE OF INVENTION: MID 4460, A HUMAN TYROSINE PHOSPHATASE
;  TITLE OF INVENTION:  FAMILY MEMBER  AND USES THEREFOR
;  FILE REFERENCE: MPI01-155P1RM
;  CURRENT APPLICATION NUMBER: US/10/245,539
;  CURRENT FILING DATE:  2002-12-05
;  PRIOR APPLICATION NUMBER: 60/323,018
;  PRIOR FILING DATE: 2001-09-18
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1118
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-245-539-2

  Query Match             16.1%;  Score 1130;  DB 14;  Length 1118;
  Best Local Similarity   25.9%;  Pred. No. 1.5e-60;
  Matches  354;  Conservative  181;  Mismatches  453;  Indels  380;  Gaps   40;

Qy          8 ARLPPRSPG-------------LRWALPLLLLLLLRLGQILCAGGTPSPIPDPSVATVAT 53

Db         25 ARAPAPNPGRNLTVETQTTSSISLSWEVP-------------DGLDS--QNSNYWVQCT 68

Qy         54 GENGITQISSTAESFHKQNGTGTPQVETNTSEDGESSGA--NDSLRTPEQGSNGTDGASQ 111

Db         69 GDGGTTETRNT-------------TATNVTVDGLGPGSLYTCSVWVEKDGVNSSVGT-- 112

Qy        112 KTPSSTGPSPVFDIKAVSISPTNVILTWKSNDTAASEYK-YVVKHKMENEK--TITVVHQ 168

Db        113 -VTTATAPNPVRNLRVEAQTNSSIALTWEVPDGPDPQNSTYGVEYTGDGGRAGTRSTAH- 170

Qy        169 PWCNIT--GLRPATSYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALS 226

Db        171 --TNITVDGLEPGCLYAFSM---------------------------------------- 188

Qy        227 WSNGNGTASCRVLLESIGSHEELTQ----DSRLQVNISDLKPGVQYNINPYLLQSNKTKG 282

Db        189 WVGKNGINSSRETRNATTAHNPVRKPESGGSDHQLHLPEL-GGPRWH-RPTELDLLRTSA 246

Qy        283 DPLGTEGGLDASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYN 342

Db        247 LEMVAEQRLETQQTPES----------------PVD--------------GLGPGSLYT 275

Qy        343 ATVYSQAANGTEGQPQAIEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIH 402

Db        276 CSVWVEKDGVNSSSWRLVTSTTAPNPVRNLTVEAQTNSSIALTWEVPDGPDPQNSTYGVE 335

Qy        403 VAGETDSSNLNVSEPRAVIPGLRSSTFYNITVCPVLGDIEGTPGFLQVH---------- 451

Db        336 YTGDGGRA------------GTRSTAHTNITVDRL------EPGCLYVFSVWVGKNGINS 377

Qy        452 --------TPPVPVSDFRVTVVSTTEIGLAWSSHDAE-----SFQMHITQEGAGNSRVEI 498

Db        378 SRETRNATTAPNPVRNLHMETQTNSSIALCWEVPDGPYPQDYTYWVGYTGDG-GGTETRN 436

Qy        499 TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHVVYVTTTEMWL 558

Db        437 TTNTSVTAERLEPGTLYTFSVWAE-KNGARGSRQNVSISTVPNAVTSLSKQDWTNSTIAL 495

Qy        559 DWKSPD--GASEYVYHL--VIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHVW 614

Db        496 RWTAPQGPGQSSYSYWVSWVREGMTDPRTQSTSGTDITLKELEAGSLYHLTVWAERNEVR 555

Qy        615 GDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNF--DDASPTYSYCLLIEKAGN------ 666

Db        556 GYNSTLTAATAPNEVTDLQNETQTKNSVMLWWKAPGDPHSQLYVYWVQWASKGHPRRGQD 615

Qy        667 -SSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFC--TDPASMA 723

Db        616 PQANWVNQTSRTNETWYKVEALEPGTLYNFTVWAERNDVASSTQ----SLCASTYPDTVT 671

Qy        724 SFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEVTYL 783

Db        672 ITSCVSTSAGYGVNLIWSCPQGGYEAFELEV-------------GGQRGSQDRS----- 712

Qy        784 NFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGFEAS 843

Db        713 ------------SCGEAVS------VLGL------------------------------ 723

Qy        844 HGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLK 903

Db        724 -GPARSYPATITT-------------IWDGMK------------------------VVS 744

Qy        904 YEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRYSDA 963

Db        745 HSVVCHTESA---------------------------------------------------- 754

Qy        964 VSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL--- 1020

Db        755 -------GVIAGAFVGIL---LFLILVGLLIFFLKRRNKKK-------QQKPELRDLVFS 797

Qy       1021 ----IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLP 1076
```

```
Db        798 SPGDIPAEDFADHVRKNERDSNCGFADEYQQLSLVGHSQSQMVASASENNAKNRYRNVLP 857

Qy       1077 YDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1135

Db        858 YDWSRVPLKPIHEEPGSDYINASFMPGLWSPQEFIATQGPLPQTVGDFWRLVWEQQSHTL 917

Qy       1136 IMLTKCVEQGRTKCEEYWP-SKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHP 1194

Db        918 VMLTNCMEAGRVKCEHYWPLDSQPCTHGHLRVTLVGEEVMENWTVRELLLLQVEEQKTLS 977

Qy       1195 LRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDR 1254

Db        978 VRQFHYQAWPDHGVPSSPDTLLAFWRMLRQWLDQTMEGGPPIVHCSAGVGRTGTLIALDV 1037

Qy       1255 LIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQ 1302

Db       1038 LLRQLQSEGLLGPFSFVRKMRESRPLMVQTEAQYVFLHQCICGSSNSQ 1085
```

RESULT 14
US-10-245-539-8
; Sequence 8, Application US/10245539
; Publication No. US20030077638A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; TITLE OF INVENTION: MID 4460, A HUMAN TYROSINE PHOSPHATASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREFOR
; FILE REFERENCE: MPI01-155P1RM
; CURRENT APPLICATION NUMBER: US/10/245,539
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: 60/323,018
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-245-539-8

```
  Query Match            16.1%;  Score 1130;  DB 14;  Length 1118;
  Best Local Similarity  25.9%;  Pred. No. 1.5e-60;
  Matches  354;  Conservative  181;  Mismatches  453;  Indels  380;  Gaps   40;

Qy          8 ARLPPRSPG-------------LRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVAT 53

Db         25 ARAPAPNPGRNLTVETQTTSSISLSWEVP-------------DGLDS--QNSNYWVQCT 68

Qy         54 GENGITQISSTAESFHKQNGTGTPQVETNTSEDGESSGA--NDSLRTPEQGSNGTDGASQ 111

Db         69 GDGGTTETRNT-------------TATNVTVDGLGPGSLYTCSVWVEKDGVNSSVGT-- 112

Qy        112 KTPSSTGPSPVFDIKAVSISPTNVILTWKSNDTAASEYK-YVVKHKMENEK--TITVVHQ 168

Db        113 -VTTATAPNPVRNLRVEAQTNSSIALTWEVPDGPDPQNSTYGVEYTGDGGRAGTRSTAH- 170

Qy        169 PWCNIT--GLRPATSYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALS 226

Db        171 --TNITVDGLEPGCLYAFSM--------------------------------------- 188

Qy        227 WSNGNGTASCRVLLESIGSHEELTQ----DSRLQVNISDLKPGVQYNINPYLLQSNKTKG 282

Db        189 WVGKNGINSSRETRNATTAHNPVRKPESGGSDHQLHLPEL-GGPRWH-RPTELDLLRTSA 246

Qy        283 DPLGTEGGLDASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYN 342

Db        247 LEMVAEQRLETQQTPES---------------PVD--------------GLGPGSLYT 275

Qy        343 ATVYSQAANGTEGQPQAIEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIH 402

Db        276 CSVWVEKDGVNSSSWRLVTSTTAPNPVRNLTVEAQTNSSIALTWEVPDGPDPQNSTYGVE 335

Qy        403 VAGETDSSNLNVSEPRAVIPGLRSSTFYNITVCPVLGDIEGTPGFLQVH----------- 451

Db        336 YTGDGGRA-----------GTRSTAHTNITVDRL------EPGCLYVFSVWVGKNGINS 377

Qy        452 --------TPPVPVSDFRVTVVSTTEIGLAWSSHDAE-----SFQMHITQEGAGNSRVEI 498

Db        378 SRETRNATTAPNPVRNLHMETQTNSSIALCWEVPDGPYPQDYTYWVGYTGDG-GGTETRN 436

Qy        499 TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHVVYVTTTEMWL 558

Db        437 TTNTSVTAERLEPGTLYTFSVWAE-KNGARGSRQNVSISTVPNAVTSLSKQDWTNSTIAL 495

Qy        559 DWKSPD--GASEYVYHL--VIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHVW 614

Db        496 RWTAPQGPGQSSYSYWVSWVREGMTDPRTQSTSGTDITLKELEAGSLYHLTVWAERNEVR 555

Qy        615 GDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNF--DDASPTYSYCLLIEKAGN------ 666

Db        556 GYNSTLTAATAPNEVTDLQNETQTKNSVMLWWKAPGDPHSQLYVYWVQWASKGHPRRGQD 615

Qy        667 -SSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFC--TDPASMA 723

Db        616 PQANWVNQTSRTNETWYKVEALEPGTLYNFTVWAERNDVASSTQ----SLCASTYPDTVT 671

Qy        724 SFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEVTYL 783

Db        672 ITSCVSTSAGYGVNLIWSCPQGGYEAFELEV-------------GGQRGSQDRS----- 712

Qy        784 NFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGFEAS 843

Db        713 ------------SCGEAVS------VLGL----------------------------- 723

Qy        844 HGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLK 903

Db        724 -GPARSYPATITT------------IWDGMK-----------------------VVS 744

Qy        904 YEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRYSDA 963

Db        745 HSVVCHTESA------------------------------------------------ 754

Qy        964 VSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL--- 1020

Db        755 -------GVIAGAFVGIL---LFLILVGLLIFFLKRRNKKK-------QQKPELRDLVFS 797

Qy       1021 ----IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLP 1076

Db        798 SPGDIPAEDFADHVRKNERDSNCGFADEYQQLSLVGHSQSQMVASASENNAKNRYRNVLP 857

Qy       1077 YDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1135

Db        858 YDWSRVPLKPIHEEPGSDYINASFMPGLWSPQEFIATQGPLPQTVGDFWRLVWEQQSHTL 917

Qy       1136 IMLTKCVEQGRTKCEEYWP-SKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHP 1194

Db        918 VMLTNCMEAGRVKCEHYWPLDSQPCTHGHLRVTLVGEEVMENWTVRELLLLQVEEQKTLS 977

Qy       1195 LRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDR 1254

Db        978 VRQFHYQAWPDHGVPSSPDTLLAFWRMLRQWLDQTMEGGPPIVHCSAGVGRTGTLIALDV 1037

Qy       1255 LIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQ 1302

Db       1038 LLRQLQSEGLLGPFSFVRKMRESRPLMVQTEAQYVFLHQCICGSSNSQ 1085
```

RESULT 15
US-10-245-539-4
; Sequence 4, Application US/10245539
; Publication No. US20030077638A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; TITLE OF INVENTION: MID 4460, A HUMAN TYROSINE PHOSPHATASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREFOR
; FILE REFERENCE: MPI01-155P1RM

```
;  CURRENT APPLICATION NUMBER: US/10/245,539
;  CURRENT FILING DATE:  2002-12-05
;  PRIOR APPLICATION NUMBER: 60/323,018
;  PRIOR FILING DATE: 2001-09-18
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 1093
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: consensus sequence
US-10-245-539-4

  Query Match             16.0%;  Score 1127;  DB 14;  Length 1093;
  Best Local Similarity   26.3%;  Pred. No. 2.2e-60;
  Matches  344;  Conservative  177;  Mismatches  438;  Indels  350;  Gaps   37;

Qy         53 TGENGITQISSTAESFHKQNGTGTPQVETNTSEDGESSGA--NDSLRTPEQGSNGTDGAS 110
              ||: | |:  :|                 || : ||   |:    |:   : | | : |
Db         43 TGDGGTTETRNT-------------TATNVTVDGLGPGSLYTCSVWVEKDGVNSSVGT- 87

Qy        111 QKTPSSTGPSPVFDIKAVSISPTNVILTWKSNDTAASEYK-YVVKHKMENEK--TITVVH 167
                ::| |:|| :::  : : ::: |||:  |    :   | |::  :  :  | :  |
Db         88 --VTTATAPNPVRNLRVEAQTNSSIALTWEVPDGPDPQNSTYGVEYTGDGGRAGTRSTAH 145

Qy        168 QPWCNIT--GLRPATSYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAAL 225
                 |||  || |   | ||:
Db        146 ---TNITVDGLEPGCLYAFSM---------------------------------------- 163

Qy        226 SWSNGNGTASCRVLLESIGSHEELTQ----DSRLQVNISDLKPGVQYNINPYLLQSNKTK 281
               |   ||  | |    :  :|  : :     |  |::: :|  | :::  |  |   :|
Db        164 -WVGKNGINSSRETRNATTAHNPVRKPESGGSDHQLHLPEL-GGPRWH-RPTELDLLRTS 220

Qy        282 GDPLGTEGGLDASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRY 341
              :  |  |:    |  |              |||                || ||: |
Db        221 ALEMVAEQRLETQQTPES---------------PVD--------------GLGPGSLY 249

Qy        342 NATVYSQAANGTEGQPQAIEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKI 401
               :|: :         : :   |    | ::|    : :|: | |:| |     | || :
Db        250 TCSVWVEKDGVNSSSWRLVTSTTAPNPVRNLTVEAQTNSSIALTWEVPDGPDPQNSTYGV 309

Qy        402 HVAGETDSSNLNVSEPRAVIPGLRSSTFYNITVCPVLGDIEGTPGFLQVH---------- 451
               |:   :             | ||:  ||||  :       || | |
Db        310 EYTGDGGRA-----------GTRSTAHTNITVDRL------EPGCLYVFSVWVGKNGIN 351

Qy        452 ---------TPPVPVSDFRVTVVSTTEIGLAWSSHDAE-----SFQMHITQEGAGNSRVE 497
                      | | || :  :   : : | | |   |         :: :  | :| | :
Db        352 SSRETRNATTAPNPVRNLHMETQTNSSIALCWEVPDGPYPQDYTYWVGYTGDG-GGTETR 410

Qy        498 ITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHVVYVTTTEMW 557
              |||  |:    | ||| | | :  :  ||  |: : |   |||:||  :     | : :
Db        411 NTTNTSVTAERLEPGTLYTFSVWAE-KNGARGSRQNVSISTVPNAVTSLSKQDWTNSTIA 469

Qy        558 LDWKSPD--GASEYVYHL--VIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHV 613
              | | :|   | | | | :  | |          ||    |||: |  |:||::|:  | : |
Db        470 LRWTAPQGPGQSSYSYWVSWVREGMTDPRTQSTSGTDITLKELEAGSLYHLTVWAERNEV 529

Qy        614 WGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNF--DDASPTYSYCLLIEKAGN----- 666
               |  ::    | |: |:::   | |  : : |     |  |  | | :       |:
Db        530 RGYNSTLTAATAPNEVTDLQNETQTKNSVMLWWKAPGDPHSQLYVYWVQWASKGHPRRGQ 589

Qy        667 --SSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFC--TDPASM 722
                :|     :    |   |  | ||: |   ::|:  |   | :    | |  | | ::
Db        590 DPQANWVNQTSRTNETWYKVEALEPGTLYNFTVWAERNDVASSTQ----SLCASTYPDTV 645

Qy        723 ASFDCEVVPKEPALVLKWTCPPGANAGPELEVSSGAWNNATHLESCSSENGTEYRTEVTY 782
               |        : | |:|| |    |||||                : |:: |:
Db        646 TITSCVSTSAGYGVNLIWSCPQGGYEAFELEV-------------GGQRGSQDRS---- 687

Qy        783 LNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGFEA 842
                        |||:  :        |:
Db        688 -------------SCGEAVS------VLGL---------------------------- 698
```

```
Qy        843 SHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVL 902
                || ::|   :||                 :|  |                    |:
Db        699 --GPARSYPATITT------------IWDGMK------------------------VV 718

Qy        903 KYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRYSD 962
                : :   ||
Db        719 SHSVVCHTESA---------------------------------------------- 729

Qy        963 AVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL-- 1020
                      ||| ||  | :   | :: ||  ||: |:|   |       | ||:   |
Db        730 --------GVIAGAFVGIL---LFLILVGLLIFFLKRRNKKK-------QQKPELRDLVF 771

Qy       1021 -----IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVL 1075
                   |  |:|  : :| : |||||||:||: | ||| || :  |  :||  |||| |||
Db        772 SSPGDIPAEDFADHVRKNERDSNCGFADEYQQLSLVGHSQSQMVASASENNAKNRYRNVL 831

Qy       1076 PYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYA 1134
                |||  ||| |  :      |||||::|||  | ::|||||||||  |: ||||:|||:  :
Db        832 PYDWSRVPLKPIHEEPGSDYINASFMPGLWSPQEFIATQGPLPQTVGDFWRLVWEQQSHT 891

Qy       1135 IIMLTKCVEQGRTKCEEYWP-SKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESH 1193
                ::||| |:| || ||| |||    |   :| : | :  | |:  ||:|:  :  ::  ::
Db        892 LVMLTNCMEAGRVKCEHYWPLDSQPCTHGHLRVTLVGEEVMENWTVRELLLLQVEEQKTL 951

Qy       1194 PLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAID 1253
                :|||: :|||||| : | |: |  ::| :: |:    | :|||||||||||| ||:|
Db        952 SVRQFHYQAWPDHGVPSSPDTLLAFWRMLRQWLDQTMEGGPPIVHCSAGVGRTGTLIALD 1011

Qy       1254 RLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQ 1302
                |: |:::|  :  :  |  :|  |||||||| |||||:||:     ||
Db       1012 VLLRQLQSEGLLGPFSFVRKMRESRPLMVQTEAQYVFLHQCICGSSNSQ 1060
```

Search completed: May 10, 2005, 19:09:31
Job time : 189 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 10, 2005, 18:54:46 ; Search time 53 Seconds
(without alignments)
2427.204 Million cell updates/sec

Title: US-09-516-728A-4
Perfect score: 7030
Sequence: 1 MKPAAREARLPPRSPGLRWA.....TIYENLAPVTTFGKTNGYIA 1337

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6826	97.1	1337	1	I38670	protein-tyrosine-p
2	4508.5	64.1	1238	2	S68700	HPTP beta-like tyr
3	1242	17.7	1997	1	S12050	protein-tyrosine-p
4	1130	16.1	1118	1	A49724	protein-tyrosine-p
5	1028	14.6	1630	2	C41214	protein-tyrosine-p
6	1026	14.6	1557	2	D41214	protein-tyrosine-p
7	981.5	14.0	583	2	S17671	protein-tyrosine-p
8	957.5	13.6	1711	1	A55148	protein-tyrosine-p
9	954	13.6	1767	2	A49502	protein-tyrosine-p
10	949	13.5	1615	2	B49502	protein-tyrosine-p
11	879	12.5	1912	2	A56178	protein-tyrosine-p
12	864	12.3	2302	2	T14328	protein-tyrosine-p
13	854	12.1	1188	1	A57064	protein-tyrosine-p
14	842	12.0	1187	1	A53661	protein-tyrosine-p
15	839	11.9	1216	2	S60613	protein-tyrosine-p
16	834.5	11.9	1226	2	JC7503	protein-tyrosine-p
17	832	11.8	1691	1	D54689	protein-tyrosine-p
18	829	11.8	1894	2	C54689	protein-tyrosine-p
19	817	11.6	1367	2	T21913	hypothetical prote
20	807	11.5	1898	2	S46216	leukocyte antigen-
21	795.5	11.3	1863	2	S46217	protein-tyrosine-p
22	792.5	11.3	1897	1	TDHULK	leukocyte antigen-
23	787.5	11.2	1907	2	S50893	protein-tyrosine-p
24	785.5	11.2	1452	1	S17670	protein-tyrosine-p
25	771.5	11.0	1452	1	S17669	protein-tyrosine-p
26	767	10.9	1437	2	T31093	probable protein-t
27	764	10.9	1290	2	A56493	leucocyte common a
28	761	10.8	2051	2	T30938	receptor tyrosine
29	750	10.7	2029	1	TDFFLK	protein-tyrosine-p
30	742	10.6	1499	2	I50212	protein-tyrosine-p
31	739.5	10.5	1501	2	I58148	protein-tyrosine-p
32	728.5	10.4	1262	1	B48758	protein-tyrosine-p
33	728.5	10.4	1496	1	A48758	protein-tyrosine-p
34	722.5	10.3	405	2	S68250	protein-tyrosine-p
35	722.5	10.3	405	2	I49372	protein-tyrosine-p
36	704	10.0	2314	1	A46151	protein-tyrosine-p
37	693	9.9	1585	2	T19121	probable protein-t
38	688.5	9.8	1440	2	JC6312	protein-tyrosine-p
39	678.5	9.7	1457	1	A48066	protein-tyrosine-p
40	674.5	9.6	1301	1	A41622	protein-tyrosine-p
41	669	9.5	1445	1	A48148	protein-tyrosine-p
42	664	9.4	1442	1	B48148	protein-tyrosine-p
43	656.5	9.3	1422	2	T42636	protein-tyrosine-p
44	640.5	9.1	1436	2	JC5290	protein-tyrosine-p
45	636.5	9.1	1304	1	A46546	leukocyte common a

ALIGNMENTS

RESULT 1
I38670
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type J precursor - human
N;Alternate names: density enhanced phosphatase-1 (DEP-1); protein-tyrosine phosphatase
C;Species: Homo sapiens (man)
C;Date: 01-Mar-1996 #sequence_revision 08-Mar-1996 #text_change 09-Jul-2004
C;Accession: I38670; I52599
R;Ostman, A.; Yang, Q.; Tonks, N.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 9680-9684, 1994
A;Title: Expression of DEP-1, a receptor-like protein-tyrosine-phosphatase, is enhanced
A;Reference number: I38670; MUID:95024024; PMID:7937872
A;Accession: I38670
A;Molecule type: mRNA
A;Residues: 1-1337 <RES>
A;Cross-references: UNIPROT:Q12913; EMBL:U10886; NID:g558754; PID:g558755
A;Experimental source: HeLa cells
R;Honda, H.; Inazawa, J.; Nishida, J.; Yazaki, Y.; Hirai, H.
Blood 84, 4186-4194, 1994
A;Title: Molecular cloning, characterization, and chromosomal localization of a novel pr
A;Reference number: I52599; MUID:95086212; PMID:7994032
A;Accession: I52599
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-216,'LTGVRKAA',225-260,'G',262-285,'GTEGGLDASNTERSRA',302,'S',304,'TAPVHDE
A;Cross-references: GB:D37781; NID:g633072; PIDN:BAA07035.1; PID:g633073
C;Comment: Enhanced expression of this protein with increasing cell density suggests a r
C;Genetics:
A;Gene: GDB:PTPRJ
A;Cross-references: GDB:385040; OMIM:600925
A;Map position: 19q13.4-19q13.4
C;Function:
A;Description: catalyzes hydrolysis of peptidyl-phosphotyrosine to peptidyl-tyrosine and
C;Superfamily: protein-tyrosine-phosphatase, receptor type J; fibronectin type III repea
C;Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; receptor; tran
F;1-35/Domain: signal sequence #status predicted <SIG>
F;36-1337/Product: protein-tyrosine-phosphatase, receptor type J #status predicted <MAT>
F;118-197/Domain: fibronectin type III repeat homology <3FNA>
F;206-283/Domain: fibronectin type III repeat homology <3FNB>
F;284-356/Domain: fibronectin type III repeat homology #status atypical <3FNC>
F;365-445/Domain: fibronectin type III repeat homology <3FND>
F;453-530/Domain: fibronectin type III repeat homology <3FNE>
F;539-617/Domain: fibronectin type III repeat homology <3FNF>
F;720-804/Domain: fibronectin type III repeat homology <3FNG>
F;972-988/Domain: transmembrane #status predicted <TMN>
F;1065-1287/Domain: protein-tyrosine-phosphatase homology <PTP>
F;72,82,93,104,142,172,192,231,258,278,342,351,376,391,396,413,431,501,525,536,582,603,6
F;1239/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1245/Binding site: substrate phosphate (Arg) #status predicted

Query Match 97.1%; Score 6826; DB 1; Length 1337;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1303; Conservative 5; Mismatches 29; Indels 0; Gaps 0;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              ||||||||||||||||||||||||||||||||||||      :||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVAHGCEEGCSLSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              ||||||||||||||||||||||||||||||||||||||||||||||  :    :    :
Db        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLAQKVAWMPAIQREAG 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              |:|         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 QGAPPPLCMMSPFVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
```

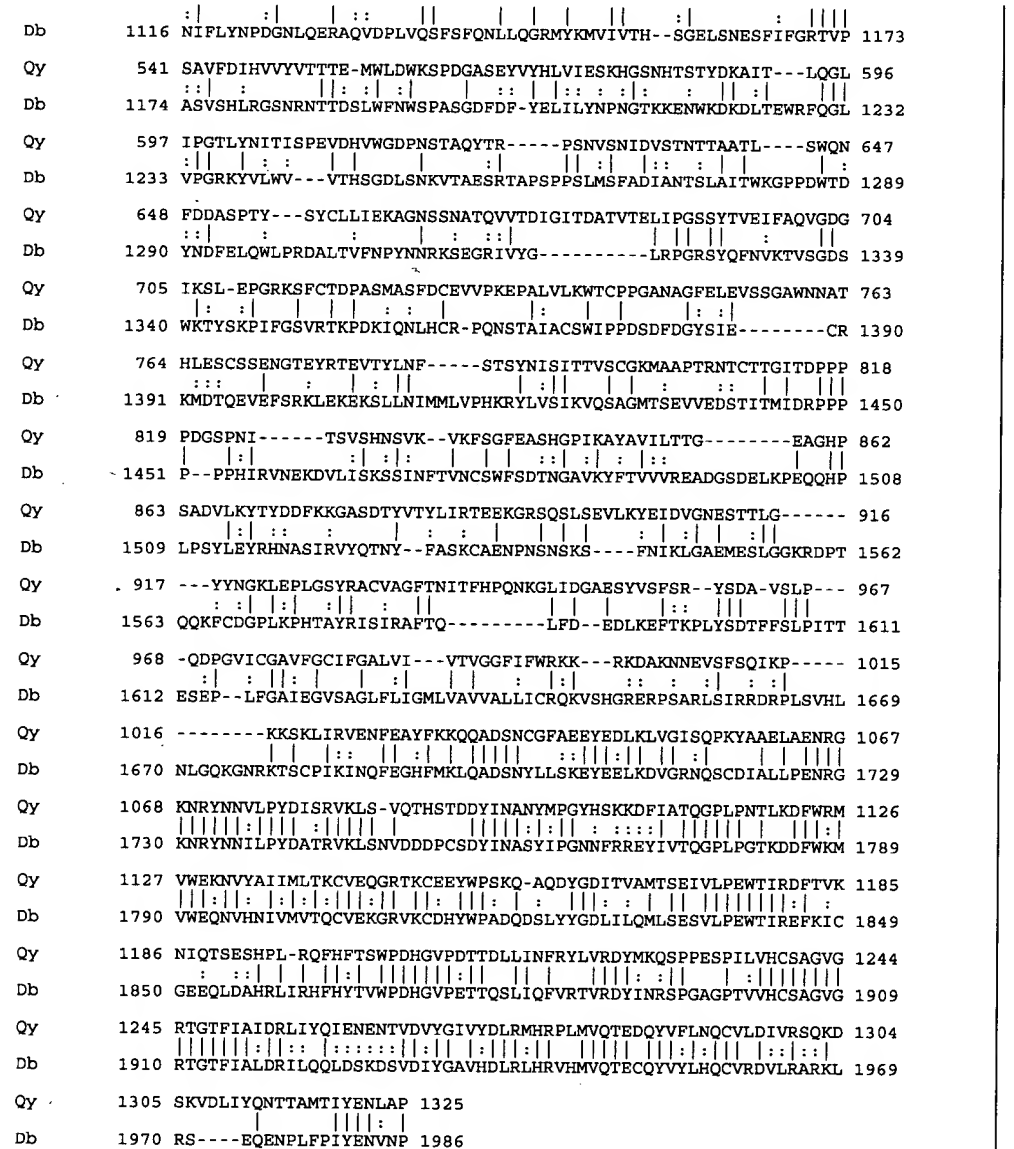
```
Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337
```

RESULT 2
S68700
HPTP beta-like tyrosine phosphatase precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S68700
R;Kuramochi, S.; Matsuda, S.; Matsuda, Y.; Saitoh, T.; Ohsugi, M.; Yamamoto, T.
FEBS Lett. 378, 7-14, 1996
A;Title: Molecular cloning and characterization of Byp, a murine receptor-type tyrosine p
A;Reference number: S68700; MUID:96140699; PMID:8549806
A;Accession: S68700
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1238 <KUR>
A;Cross-references: UNIPROT:Q64455; GB:D45212; NID:g1208432; PIDN:BAA08146.1; PID:g120843
C;Genetics:
A;Map position: 2E1-2
C;Superfamily: protein-tyrosine-phosphatase, receptor type J; fibronectin type III repeat
C;Keywords: phosphoprotein
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-1238/Product: HPTP beta-like tyrosine phosphatase #status predicted <MAT>
F;267-347/Domain: fibronectin type III repeat homology <3FR>
F;966-1188/Domain: protein-tyrosine-phosphatase homology <PTP>
F;1140/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1146/Binding site: substrate phosphate (Arg) #status predicted

Query Match 64.1%; Score 4508.5; DB 2; Length 1238;
Best Local Similarity 67.2%; Pred. No. 1.5e-244;
Matches 901; Conservative 108; Mismatches 226; Indels 105; Gaps 12;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||| | |||||||||||  |||||| ||:||||  |:||
Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLRQGQVLCAGAAPNPI---------------- 43

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||| | |||||||||||  |||||| ||:||||  |:||
Db         44 ---------------------------------------------------------- 43

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              |||:|| :|||:|:|||| ||: ||| :  :::|||:  |  | :|  ||||||  | |
Db         44 --FDIEAV-VSPTSVLLTWKHNDSGASECR--IENKMESNLTFPVKNQTSCNITGLSPGT 98

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              || |||     |||        : ||| |||||  |   || :| |:||| |||||  |:|:
Db         99 SYTFSIISVTTNETLNK----TITTEPWPVSDLHVTSVGVTQARLTWSNANGTASYRMLI 154

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLL-QSNKTKGDPLGTEGGLDASNTERS 299
              | : :|          |||| ||||     | :   :||:|: |    |   ||: |:|
Db        155 EELTTHS--------SVNISGLKPGTN---NTFAFPESNETQADFAVAEEVPDANGTKRI 203

Qy        300 RAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQA 359
                :  : :|  |||  ||| |||    ||:||  |:| |:||||:||||||||||||:
```

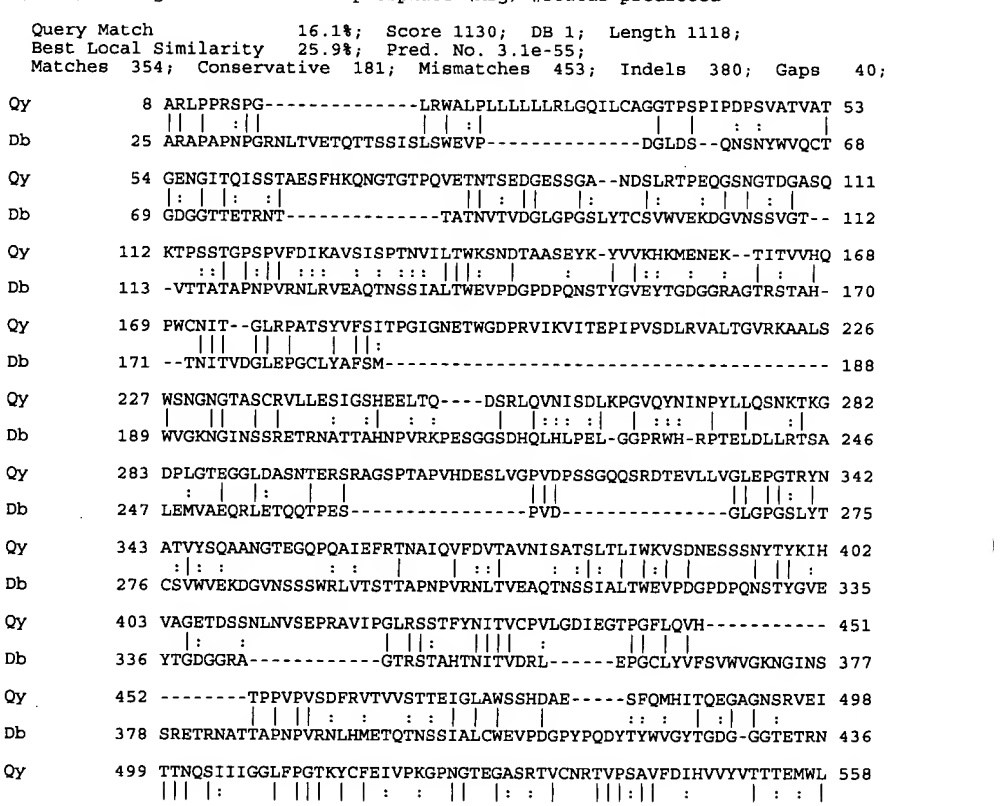
```
Db        204 PV-TNLSQLHKNSLVS-VDPPSGQDPSLTEILLTDLKPDTQYNATIYSQAANGTEGQPRN 261

Qy        360 IEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419

Db        262 KVFKTNSTQVSDVRAMNISASSMTLTWKSNYDGSRTSIVYKIHVAGGTHSVNQTVNKTEA 321

Qy        420 VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA 479

Db        322 IILGLSSSTLYNITVHPFLGQTEGTPGFLQVYTSPDQVSDFRVTNVSTRAIGLAWRSNDS 381

Qy        480 ESFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTV 539

Db        382 KSFEIFIKQDGGEKHRNASTGNQSYMVEDLKPGTSYHFEIIPRGPDGTEGLSSTVNGSTD 441

Qy        540 PSAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGS-NHTSTYDKAITLQGLIP 598

Db        442 PSAVTDIRVVNISTTEMQLEWQNTDDASGYTYHLVLESKSGSIIRTNSSQKWITVGSLTP 501

Qy        599 GTLYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYC 658

Db        502 GTLYNVTIFPEVDQIQGISNSITQYTRPSSVSHIEVNTTTTTAAIRWKNEDAASASYAYS 561

Qy        659 LLIEKAGNSSNATQVVT-DIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCT 717

Db        562 VLILKTGDGSNVTSNFTKDPSI---LIPELIPGVSYTVKILTQVGDGTTSLVPGWNLFCT 618

Qy        718 DPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYR 777

Db        619 EPEPVTSFHCEVVPKEPALVLKWACPFGMYTGFELGVRSDSWDNMTRLENCTSDDDTECR 678

Qy        778 TEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKF 837

Db        679 TEVAYLNFSTSYNISIATLSCGKMALPAQNICTTGITDPPTPDGSPNITSVSHNSVKVKF 738

Qy        838 SGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQS 897

Db        739 SGFEASHGPIKAYAVILTTGEAAQPSADVLKYTYEDFKRGASDTYVTYLIRIEEKGQSQG 798

Qy        898 LSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSF 957

Db        799 LSEVLNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYVSF 858

Qy        958 SRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKK 1017

Db        859 SPYSEAVFLPQDPGVICGAVFGCIFGALAITAVGGFIFWRKKRTDAKNNEVSFSQIKPKK 918

Qy       1018 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPY 1077

Db        919 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYTAEIAENRGKNRYNNVLPY 978

Qy       1078 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIM 1137

Db        979 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIVM 1038

Qy       1138 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQ 1197

Db       1039 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEVVLPEWTIRDFVVKNMQNSESHPLRQ 1098

Qy       1198 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIY 1257

Db       1099 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRLIY 1158

Qy       1258 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAM 1317

Db       1159 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTTAM 1218

Qy       1318 TIYENLAPVTTFGKTNGYIA 1337

Db       1219 TIYENLEPVSMFGKTNGYIA 1238


RESULT 3
S12050
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type beta precursor - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S12050; S15818; S15819
R;Krueger, N.X.; Streuli, M.; Saito, H.
EMBO J. 9, 3241-3252, 1990
A;Title: Structural diversity and evolution of human receptor-like protein tyrosine phos
A;Reference number: S12049; MUID:91006018; PMID:2170109
A;Accession: S12050
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1997 <KRU>
A;Cross-references: UNIPROT:P23467; GB:X54131; NID:g35787; PIDN:CAA38066.1; PID:g35788
R;de Vries, L.; Li, R.Y.; Ragab, A.; Ragab-Thomas, J.M.F.; Chap, H.
FEBS Lett. 282, 285-288, 1991
A;Title: Expression of a truncated protein-tyrosine phosphatase mRNA in human lung.
A;Reference number: S15818; MUID:91243813; PMID:1645282
A;Accession: S15818
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1872-1911,'VHMVLQK' <VRI>
A;Accession: S15819
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1872-1997 <VR2>
C;Genetics:
A;Gene: GDB:PTPRB; PTPB
A;Cross-references: GDB:127352; OMIM:176882
A;Map position: 12q15-12q21
C;Superfamily: protein-tyrosine-phosphatase, receptor type beta; fibronectin type III re
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-1997/Product: protein-tyrosine-phosphatase, receptor type beta #status predicted <M
F;23-1625/Domain: extracellular #status predicted <EXT>
F;1626-1642/Domain: transmembrane #status predicted <TMN>
F;1643-1997/Domain: intracellular #status predicted <INT>
F;1727-1952/Domain: protein-tyrosine-phosphatase homology <PTP>
F;1904/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1910/Binding site: substrate phosphate (Arg) #status predicted

  Query Match             17.7%;  Score 1242;  DB 1;  Length 1997;
  Best Local Similarity   29.6%;  Pred. No. 4.1e-61;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182

Db        759 FDHYEVTIKNKNNFIQTKSIP------------KSENE----------CVFVQLVPGRLY 796

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238

Db        797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy        239 LLESIGSHE--ELTQDSRL-----------QVNISDLKPGVQY-----NINPYLLQSNKT 280

Db        845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy        281 KG------------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313

Db        902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373

Db        958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424

Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484

Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
```



RESULT 4
A49724
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type H precursor - human
N;Alternate names: protein-tyrosine-phosphatase, stomach cancer-associated type 1; SAP-1
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C;Accession: A49724
R;Matozaki, T.; Suzuki, T.; Uchida, T.; Inazawa, J.; Ariyama, T.; Matsuda, K.; Horita, K
J. Biol. Chem. 269, 2075-2081, 1994

A;Title: Molecular cloning of a human transmembrane-type protein tyrosine phosphatase and
A;Reference number: A49724; MUID:94124561; PMID:8294459
A;Accession: A49724
A;Molecule type: mRNA
A;Residues: 1-1118 <MATO>
A;Cross-references: UNIPROT:Q15426; GB:D15049; NID:g475003; PIDN:BAA03645.1; PID:g475004
C;Genetics:
A;Gene: GDB:PTPRH; SAP-1
A;Cross-references: GDB:305504
A;Map position: 19q13.4-19q13.4
A;Note: highly expressed in colon and pancreatic cancer cells but not in the normal cells
C;Superfamily: protein-tyrosine-phosphatase, receptor type H; fibronectin type III repeat
C;Keywords: carcinogenesis; duplication; glycoprotein; phosphoprotein; phosphoric monoest
F;1-27/Domain: signal sequence #status predicted <SIG>
F;27-110/Domain: fibronectin type III repeat homology <3FNA>
F;28-1118/Product: protein-tyrosine-phosphatase, receptor type H #status predicted <MAT>
F;28-761/Domain: extracellular #status predicted <EXT>
F;116-199/Domain: fibronectin type III repeat homology <3FNB>
F;205-289/Domain: fibronectin type III repeat homology <3FNC>
F;296-379/Domain: fibronectin type III repeat homology <3FND>
F;385-468/Domain: fibronectin type III repeat homology <3FNE>
F;474-558/Domain: fibronectin type III repeat homology <3FNF>
F;564-658/Domain: fibronectin type III repeat homology <3FNG>
F;667-737/Domain: fibronectin type III repeat homology <3FNH>
F;762-778/Domain: transmembrane #status predicted <TMN>
F;779-1118/Domain: intracellular #status predicted <INT>
F;846-1070/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;35,78,83,107,132,149,172,196,203,286,304,312,329,352,376,383,401,436,439,470,490,558,5
F;1022/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1028/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.1%; Score 1130; DB 1; Length 1118;
Best Local Similarity 25.9%; Pred. No. 3.1e-55;
Matches 354; Conservative 181; Mismatches 453; Indels 380; Gaps 40;



```
Db        437 TTNTSVTAERLEPGTLYTFSVWAE-KNGARGSRQNVSISTVPNAVTSLSKQDWTNSTIAL 495

Qy        559 DWKSPD--GASEYVYHL--VIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHVW 614

Db        496 RWTAPQGPGQSSYSYWVSWVREGMTDPRTQSTSGTDITLKELEAGSLYHLTVWAERNEVR 555

Qy        615 GDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNF--DDASPTYSYCLLIEKAGN------ 666

Db        556 GYNSTLTAATAPNEVTDLQNETQTKNSVMLWWKAPGDPHSQLYVYWVQWASKGHPRRGQD 615

Qy        667 -SSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFC--TDPASMA 723

Db        616 PQANWVNQTSRTNETWYKVEALEPGTLYNFTVWAERNDVASSTQ----SLCASTYPDTVT 671

Qy        724 SFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEVTYL 783

Db        672 ITSCVSTSAGYGVNLIWSCPQGGYEAFELEV-------------GGQRGSQDRS----- 712

Qy        784 NFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGFEAS 843

Db        713 ------------SCGEAVS------VLGL---------------------------- 723

Qy        844 HGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLK 903

Db        724 -GPARSYPATITT-------------IWDGMK------------------------VVS 744

Qy        904 YEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRYSDA 963

Db        745 HSVVCHTESA---------------------------------------------- 754

Qy        964 VSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL--- 1020

Db        755 -------GVIAGAFVGIL---LFLILVGLLIFFLKRRNKKK-------QQKPELRDLVFS 797

Qy       1021 ----IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLP 1076

Db        798 SPGDIPAEDFADHVRKNERDSNCGFADEYQQLSLVGHSQSQMVASASENNAKNRYRNVLP 857

Qy       1077 YDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1135

Db        858 YDWSRVPLKPIHEEPGSDYINASFMPGLWSPQEFIATQGPLPQTVGDFWRLVWEQQSHTL 917

Qy       1136 IMLTKCVEQGRTKCEEYWP-SKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHP 1194

Db        918 VMLTNCMEAGRVKCEHYWPLDSQPCTHGHLRVTLVGEEVMENWTVRELLLLQVEEQKTLS 977

Qy       1195 LRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDR 1254

Db        978 VRQFHYQAWPDHGVPSSPDTLLAFWRMLRQWLDQTMEGGPPIVHCSAGVGRTGTLIALDV 1037

Qy       1255 LIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQ 1302

Db       1038 LLRQLQSEGLLGPFSFVRKMRESRPLMVQTEAQYVFLHQCICGSSNSQ 1085


RESULT 5
C41214
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type 10D, long splice form precurso
C;Species: Drosophila melanogaster
C;Date: 28-May-1992 #sequence_revision 12-Jun-1992 #text_change 24-Apr-1998
C;Accession: C41214
R;Yang, X.; Seow, K.T.; Bahri, S.M.; Oon, S.H.; Chia, W.
Cell 67, 661-673, 1991
A;Title: Two Drosophila receptor-like tyrosine phosphatase genes are expressed in a subs
A;Reference number: A41214; MUID:92034988; PMID:1657401
A;Accession: C41214
A;Molecule type: mRNA
A;Residues: 1-1630 <YAN>
A;Cross-references: GB:M80465
C;Genetics:
A;Gene: FlyBase:Ptp10D
A;Cross-references: FlyBase:FBgn0004370
C;Superfamily: protein-tyrosine-phosphatase, receptor type 4E; fibronectin type III repe

C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept
F;1197-1213/Domain: transmembrane #status predicted <TMN>
F;1214-1630/Domain: intracellular #status predicted <INT>
F;1295-1515/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1467/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1473/Binding site: substrate phosphate (Arg) #status predicted

  Query Match           14.6%;  Score 1028;  DB 2;  Length 1630;
  Best Local Similarity  25.7%;  Pred. No. 2.9e-49;
  Matches  330;  Conservative  232;  Mismatches  470;  Indels  250;  Gaps   52;

Qy        119 PSPVFDIKAVSISPTN-VILTWKSNDTAASEYKYVVKHKME--NEKTITV-VHQPWCNIT 174

Db        405 PLPVRNLRSINDDKTNTMIITWEADPASTQDEYRIVYHELETFNGDTSTLTTDRTRFTLE 464

Qy        175 GLRPATSYVFSITPGIGNETWGDPRVIKVITEPIP--VSDLRVALTGVRKAALSWSNGNG 232

Db        465 SLLPGRNYSLSV-QAVSKKMESNETSIFVVTRPSSPIIEDLKSIRMGLN---ISWKS--- 517

Qy        233 TASCRVLLESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLD 292

Db        518 ---------DVNSKQE-----------------QYEV-------------LYSRNGTS 536

Qy        293 ASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANG 352

Db        537 DLRT-------------------------QKTKESRLVIKNLQPGAAYELKVFA-VSHD 569

Qy        353 TEGQPQAIEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNL 412

Db        570 LRSEPHAYFQAVYPNPPRNMTIETVRSNSVLVHWSPPESGEFTEYSIRYRTDSEQQWVRL 629

Qy        413 -NVSEPRAVIPGLRSSTFYNITVCPVLGDIEG-TPGFLQVHTPPVPVSDFRVTVVSTTEI 470

Db        630 PSVRSTEADITDMTKGEKYTIQVNTVSFGVESPVPQEVNTTVPPNPVSNI-IQLVDSRNI 688

Qy        471 GLAWSSHD--AESFQMH----------ITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFE 518

Db        689 TLEWPKPEGRVESYILKWWPSDNPGRVQTKNVSENKSADDLSTVRVLIGELMPGVQYKFD 748

Qy        519 IVPKGPNGTEGASRTVCNRTVPSAVFDIHVV--------------YVTTTEMWLDWKSPD 564

Db        749 IQTTSYGILSGIT-SLYPRTMPLIQSDVVVANGEKEDERDTITLSYTPT---------PQ 798

Qy        565 GASEY-VYHL------VIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHVWGDP 617

Db        799 SSSKFDIYRFSSGDAEIRDKEKLANDT---DRKVTFTGLVPGRLYNIT-------VWTVS 848

Qy        618 NSTAQ---------YTRP------SNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLLIE 662

Db        849 GGVASLPIQRQDRLYPEPITQLHATNITDTEISLRWDLPKGEYNDFDIAYLTADNLL--- 905

Qy        663 KAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVG---DGIKSLEPGRKSFCTD- 718

Db        906 -AQNMTTRNEI---------TISDLRPHRNYTFTVVVRSGTESSVLRSSSPLSASFTTNE 955

Qy        719 --PASMASF-DCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTE 775

Db        956 AVPGRVERFHPTDVQPSE--INFEWSL-PSSEANGVIRQFSIAYTNINNLTDAGMQ---D 1009

Qy        776 YRTEVTY-----LNFSTSYNISI---TTVSCGKMAAPTRN-TCTTGITDPPPPDGSPNIT 826

Db       1010 FESEEAFGVIKNLKPGETYVFKIQAKTAIGFG----PEREYRQTMPILAPPRPATQVVPT 1065

Qy        827 SV--SHNSVKVKF--SGFEASHGPIKAYAVILTTGEAGHPSA-------DVLKYTYDDFK 875

Db       1066 EVYRSSSTIQIRFRKNYFSDQNGQVRMYTIIVAEDDAKNASGLEMPSWLDVQSYSV--WL 1123

Qy        876 KGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNEST---TLGYYNGKLEPLGSYRACV 932

Db       1124 PYAIDPYYPFENRSVE------------DFTIGTENCDNHKIGYCNGPLKSGTTIGVKV 1170

Qy        933 AGFTNITFHPQNKGLIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGG 992
```

```
Db       1171 RRFT------------GADKFTD-TAYSFPIQTDQDNTSLIVAI--TVPLTIILVLLVT 1214

Qy        993 FIFWRKKR-------KDAK-NNEVSFSQIKPKKSKLIRVENFEAYFKKQQADSNCGFAEE 1044
Db       1215 LIFYKRRRNNCRKTTKDSRANDNMSLPDSVIEQNRPILIKNFAEHYRLMSADSDFRFSEE 1274

Qy       1045 YEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGY 1103
Db       1275 FEELKHVGRDQPCTFADLPCNRPKNRFTNILPYDHSRFKLQPVDDDEGSDYINANYVPGH 1334

Qy       1104 HSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQ-DYG 1162
Db       1335 NSPREFIVTQGPLHSTRDDFWRMCWESNSRAIVMLTRCFEKGREKCDQYWPNDTVPVFYG 1394

Qy       1163 DITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLV 1222
Db       1395 DIKVQILNDSHYADWVMTEFML--CRGSEQRILRHFHFTTWPDFGVPNPPQTLVRFVRAF 1452

Qy       1223 RDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMV 1282
Db       1453 RD--RICAEQRPIVVHCSAGVGRSGTFITLDRILQQINTSDYVDIFGIVYAMRKERVWMV 1510

Qy       1283 QTEDQYVFLNQCVLDIVRSQKD 1304
Db       1511 QTEQQYICIHQCLLAVLEGKEN 1532


RESULT 6
D41214
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type 10D, short splice form precurs
C;Species: Drosophila melanogaster
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 09-Jul-2004
C;Accession: D41214; A41215
R;Yang, X.; Seow, K.T.; Bahri, S.M.; Oon, S.H.; Chia, W.
Cell 67, 661-673, 1991
A;Title: Two Drosophila receptor-like tyrosine phosphatase genes are expressed in a subs
A;Reference number: A41214; MUID:92034988; PMID:1657401
A;Accession: D41214
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1557 <YAN>
A;Cross-references: UNIPROT:P35992; GB:M80465
R;Tian, S.S.; Tsoulfas, P.; Zinn, K.
Cell 67, 675-685, 1991
A;Title: Three receptor-linked protein-tyrosine phosphatases are selectively expressed o
A;Reference number: A41215; MUID:92034989; PMID:1657402
A;Accession: A41215
A;Molecule type: mRNA
A;Residues: 1-904,'L',906-1125,'Q',1126-1165,'YR',1168-1171,'A',1173-1215,'L',1217-1456,
A;Cross-references: GB:M80538; NID:g158644; PIDN:AAA28952.1; PID:g158645
C;Genetics:
A;Gene: FlyBase:Ptp10D
A;Cross-references: FlyBase:FBgn0004370
C;Superfamily: protein-tyrosine-phosphatase, receptor type 4E; fibronectin type III repe
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept
F;1197-1213/Domain: transmembrane #status predicted <TMN>
F;1214-1557/Domain: intracellular #status predicted <INT>
F;1295-1515/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1467/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1473/Binding site: substrate phosphate (Arg) #status predicted

  Query Match           14.6%;  Score 1026;  DB 2;  Length 1557;
  Best Local Similarity  25.7%;  Pred. No. 3.5e-49;
  Matches  329;  Conservative  233;  Mismatches  470;  Indels  250;  Gaps   52;

Qy        119 PSPVFDIKAVSISPTN-VILTWKSNDTAASEYKYVVKHKME--NEKTITV-VHQPWCNIT 174
Db        405 PLPVRNLRSINDDKTNTMIITWEADPASTQDEYRIVYHELETFNGDTSTLTTDRTRFTLE 464

Qy        175 GLRPATSYVFSITPGIGNETWGDPRVIKVITEPIP--VSDLRVALTGVRKAALSWSNGNG 232
Db        465 SLLPGRNYSLSV-QAVSKKMESNETSIFVVTRPSSPIIEDLKSIRMGLN---ISWKS--- 517

Qy        233 TASCRVLLESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLD 292
Db        518 ---------DVNSKQE-----------------QYEV-------------LYSRNGTS 536

Qy        293 ASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANG 352
Db        537 DLRT-------------------------QKTKESRLVIKNLQPGAAYELKVFA-VSHD 569

Qy        353 TEGQPQAIEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNL 412
Db        570 LRSEPHAYFQAVYPNPPRNMTIETVRSNSVLVHWSPPESGEFTEYSIRYRTDSEQQWVRL 629

Qy        413 -NVSEPRAVIPGLRSSTFYNITVCPVLGDIEG-TPGFLQVHTPPVPVSDFRVTVVSTTEI 470
Db        630 PSVRSTEADITDMTKGEKYTIQVNTVSFGVESPVPQEVNTTVPPNPVSNI-IQLVDSRNI 688

Qy        471 GLAWSSHD--AESFQMH----------ITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFE 518
Db        689 TLEWPKPEGRVESYILKWWPSDNPGRVQTKNVSENKSADDLSTVRVLIGELMPGVQYKFD 748

Qy        519 IVPKGPNGTEGASRTVCNRTVPSAVFDIHVV--------------YVTTTEMWLDWKSPD 564
Db        749 IQTTSYGILSGIT-SLYPRTMPLIQSDVVVANGEKEDERDTITLSYTPT---------PQ 798

Qy        565 GASEY-VYHL------VIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHVWGDP 617
Db        799 SSSKFDIYRFSSGDAEIRDKEKLANDT---DRKVTFTGLVPGRLYNIT-------VWTVS 848

Qy        618 NSTAQ---------YTRP------SNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLLIE 662
Db        849 GGVASLPIQRQDRLYPEPITQLHATNITDTEISLRWDLPKGEYNDFDIAYLTADNLI--- 905

Qy        663 KAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVG---DGIKSLEPGRKSFCTD- 718
Db        906 -AQNMTTRNEI---------TISDLRPHRNYTFTVVVRSGTESSVLRSSSPLSASFTTNE 955

Qy        719 --PASMASF-DCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTE 775
Db        956 AVPGRVERFHPTDVQPSE--INFEWSL-PSSEANGVIRQFSIAYTNINNLTDAGMQ---D 1009

Qy        776 YRTEVTY-----LNFSTSYNISI---TTVSCGKMAAPTRN-TCTTGITDPPPPDGSPNIT 826
Db       1010 FESEEAFGVIKNLKPGETYVFKIQAKTAIGFG----PEREYRQTMPILAPPRPATQVVPT 1065

Qy        827 SV--SHNSVKVKF--SGFEASHGPIKAYAVILTTGEAGHPSA-------DVLKYTYDDFK 875
Db       1066 EVYRSSSTIQIRFRKNYFSDQNGQVRMYTIIVAEDDAKNASGLEMPSWLDVQSYSV--WL 1123

Qy        876 KGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNEST---TLGYYNGKLEPLGSYRACV 932
Db       1124 PYAIDPYYPFENRSVE------------DFTIGTENCDNHKIGYCNGPLKSGTTIGVKV 1170

Qy        933 AGFTNITFHPQNKGLIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGG 992
Db       1171 RRFT------------GADKFTD-TAYSFPIQTDQDNTSLIVAI--TVPLTIILVLLVT 1214

Qy        993 FIFWRKKR-------KDAK-NNEVSFSQIKPKKSKLIRVENFEAYFKKQQADSNCGFAEE 1044
Db       1215 LIFYKRRRNNCRKTTKDSRANDNMSLPDSVIEQNRPILIKNFAEHYRLMSADSDFRFSEE 1274

Qy       1045 YEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGY 1103
Db       1275 FEELKHVGRDQPCTFADLPCNRPKNRFTNILPYDHSRFKLQPVDDDEGSDYINANYVPGH 1334

Qy       1104 HSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQ-DYG 1162
Db       1335 NSPREFIVTQGPLHSTRDDFWRMCWESNSRAIVMLTRCFEKGREKCDQYWPNDTVPVFYG 1394

Qy       1163 DITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLV 1222
Db       1395 DIKVQILNDSHYADWVMTEFML--CRGSEQRILRHFHFTTWPDFGVPNPPQTLVRFVRAF 1452

Qy       1223 RDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMV 1282
```

```
                ||  :    : ||:||||||||||:|||| :||:: ||   : ||::|||| :|  |  ||
Db       1453 RD--RICAEQRPIVVHCSAGVGRSGTFITLDRILQQINTSDYVDIFGIVYAMRKERVWMV 1510

Qy       1283 QTEDQYVFLNQCVLDIVRSQKD 1304
                ||| ||: ::||:| ::  :::
Db       1511 QTEQQYICIHQCLLAVLEGKEN 1532


RESULT 7
S17671
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type beta - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 15-Mar-1996 #text_change 09-Jul-2004
C;Accession: S17671; S40287
R;Gebbink, M.F.B.G.; van Etten, I.; Hateboer, G.; Suijkerbuijk, R.; Beijersbergen, R.L.;
FEBS Lett. 290, 123-130, 1991
A;Title: Cloning, expression and chromosomal localization of a new putative receptor-lik
A;Reference number: S17669; MUID:92008644; PMID:1655529
A;Accession: S17671
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-583 <GEB>
A;Cross-references: UNIPROT:Q64497; EMBL:X58289
R;Hendriks, W.; Brugman, C.; Zeeuwen, P.; Schepens, J.; Wieringa, B.
submitted to the EMBL Data Library, June 1993
A;Description: Assessment of the expression levels of murine protein-tyrosine phosphatas
A;Reference number: S40280
A;Accession: S40287
A;Molecule type: mRNA
A;Residues: 377-483,'T',485-486 <HEN>
A;Cross-references: EMBL:Z23056; NID:g438149; PIDN:CAA80591.1; PID:g438150
C;Superfamily: protein-tyrosine-phosphatase, receptor type beta; fibronectin type III re
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane prot
F;311-536/Domain: protein-tyrosine-phosphatase homology <PTP>
F;488/Active site: Cys (phosphocysteine intermediate) #status predicted
F;494/Binding site: substrate phosphate (Arg) #status predicted

  Query Match             14.0%;  Score 981.5;  DB 2;  Length 583;
  Best Local Similarity   39.0%;  Pred. No. 2.4e-47;
  Matches  226;  Conservative  102;  Mismatches  191;  Indels   61;  Gaps   15;

Qy        789 YNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNI------TSVSHNSVK--VKFSGF 840
              | :||   | |  :    ::  | |  || |   |:|        :| :|:   |  | |
Db          5 YLVSIKVQSAGMTSEVVEDSTITMIDRPPQP--PPHIRVNEKDVLISKSSINFTVNCSWF 62

Qy        841 EASHGPIKAYAVILTTG--------EAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEK 892
              ::| :  :||::             |  ||    |:| ::   :      |  :  :  |
Db         63 SDTNGAVGYFAVVVREADSMDELKPEQQHPLPSYLEYRHNASIRVYQTNY--FASKCAES 120

Qy        893 GRSQSLSEVLKYEIDVGNESTTLG---------YYNGKLEPLGSYRACVAGFTNITFHPQ 943
              | | |    : | :| |  :||         : :| | |  :||  :  || :
Db        121 PDSSSKS----FNIKLGAEMDSLGGKCDPSQQKFCDGPLLPHTAYRISIRAFTQLF---- 172

Qy        944 NKGLIDGAESYVSFSRYSDAVSLPQDP--GVICGAVFGCIFGALVIVTVGGFIFWRK--K 999
              :: | :  :   | : :|  ::   :|  ||| |   |     :::  |  ||  :|
Db        173 DEDLKEFTKPLYSDTFFSMPITTESEPLFGVIEGVSAGLFLIGMLVALVAFFICRQKASH 232

Qy       1000 RKDAKNNEVSFSQIKP------------KKSKLIRVENFEAYFKKQQADSNCGFAEEYE 1046
              ::  :  :|  : :|              | |  |::  || :| | |||||   ::|||
Db        233 SRERPSARLSIRRDRPLSVHLNLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYE 292

Qy       1047 DLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHS 1105
              ||| || ||    | | ||||||||||:|||| ||||||  |       |||||:|:|| :
Db        293 DLKDVGRSQSCDIALLPENRGKNRYNNILPYDASRVKLSNVDDDPCSDYINASYIPGNNF 352

Qy       1106 KKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDI 1164
              ::::|||||||| |  |||:| ||:||: |:|:|:|:|||:|| ||: |||: |    |||:
Db        353 RREYIATQGPLPGTKDDFWKMAWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDPLYYGDL 412

Qy       1165 TVAMTSEIVLPEWTIRDFTVKNIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVR 1223
              : | || |||||||||:| : : :  ::|  | | ||:| |||||||:||  || |   ||
Db        413 ILQMVSESVLPEWTIREFKICSEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVR 472

Qy       1224 DYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQ 1283
              ||: :||   | :|||||||||||:|:||:: |:::::||:|| |:|||:||  |||
Db        473 DYINRSPGAGPSVVHCSAGVGRTGTFVALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQ 532

Qy       1284 TEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIYENL 1323
              || |||:|:||| |:   :         |    ::||||:
Db        533 TECQYVYLHQCVRDVSEQRNCGN----EQEKGGVSIYENV 568


RESULT 8
A55148
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type OST precursor - rat
N;Alternate names: OST-PTP; osteotesticular protein-tyrosine-phosphatase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A55148
R;Mauro, L.J.; Olmsted, E.A.; Skrobacz, B.M.; Mourey, R.J.; Davis, A.R.; Dixon, J.E.
J. Biol. Chem. 269, 30659-30667, 1994
A;Title: Identification of a hormonally regulated protein tyrosine phosphatase associate
A;Reference number: A55148; MUID:95074080; PMID:7527035
A;Accession: A55148
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1711 <MAU>
A;Cross-references: GB:L36884
C;Comment: The sequence contains ten fibronectin type III repeats and two protein-tyrosi
C;Superfamily: protein-tyrosine-phosphatase, receptor type OST; fibronectin type III rep
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane prot
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1711/Product: protein-tyrosine-phosphatase, receptor type OST #status predicted <MA
F;1174-1398/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1350/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1356/Binding site: substrate phosphate (Arg) #status predicted

  Query Match             13.6%;  Score 957.5;  DB 1;  Length 1711;
  Best Local Similarity   24.5%;  Pred. No. 2.8e-45;
  Matches  383;  Conservative  227;  Mismatches  524;  Indels  431;  Gaps   72;

Qy         22 PLLLL--LLRLGQIL-----CA--GGTP-----SPIPDPSVAT----------VATGEN 56
              ||:||  || |   |     |:  ||:|     |:   :|::           ||
Db          3 PLILLAALLWLQGFLAEDDACSSLGGSPDRQGGGPLLSVNVSSHGKSTSLFLSWVAAELG 62

Qy         57 GITQISSTAESFHKQNGTGTP---QVETNTSEDG-ESSGANDS---------LRTPEQGS 103
              |       | |    : :|:|    |:: :|:| | |  |               ||   |
Db         63 GF----DYALSLRSVDSSGSPEGQQLQAHTNESGFEFHGLVPGSRYQLKLTVLRPCWQNV 118

Qy        104 NGTDGASQKTPSSTGPSPVFDIKAVSI-SPTNVILTWKSNDTAASEYKYVVKHKMENEKT 162
              |       : | |: |  ::  |  ||  :  :|          |: :: | :|:: |
Db        119 TIT------LTARTAPTVVRGLQLHSAGSPARLEASWSDAPGDQDSYQLLLYH-LESQ-T 170

Qy        163 ITVVHQPWCNIT----------GLRPATSYVFSITPGIGNETWGDPRVIKVI----TEP 207
              :       ||::          | | | ||  :       ||      |       |||
Db        171 LA------CNVSVSPDTLSYSFGDLLPGTQYVLEVI------TWAGSLHAKTSILQWTEP 218

Qy        208 IPVSDLRVALTGVRKAALSWSNGNGTASCRVLLESIGSHEELTQDSRLQVN---ISDLKP 264
              :|    | :   |       |::  |  |  ::|  :       |    |  |:      | |
Db        219 VPPDHLALRALGTSSLQAFWNSSEGATSFHLMLTDLLGGTNTTAVIRQGVSTHTFLHLSP 278

Qy        265 GVQYNINPYLLQSNKTKG-----DPLGTE------------------------------ 288
              |       |: |:   : |      |  ||
Db        279 G-----TPHELKICASAGPHQIWGPSATEWTYPSYPSDLVLTPLRNELWASWKAGLGARD 333

Qy        289 -------GGLDASNT---ERSRA--GSPTAPVHD----ESLVGPVD---------PSSGQ 323
                    | :::::|   |   |     |  | |      : | || |           |
Db        334 GYVLKLSGPMESTSTLGPEECNAVFPGPLPPGHYTLQLKVLAGPYDAWVEGSTWLAESAA 393

Qy        324 QSRD---TEVLLVGLE----PGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVTAVN 376
              |:     : | |||     || |  | :||  | |: |                    |
Db        394 LPREVPGARLWLDGLEASKQPGRR--ALLYSDDAPGSLG-------------------N 431

Qy        377 ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAVIPGLRSSTFYNITVCP 436
              ||  |                  | |              :  ||     | : :
Db        432 ISVPS----------------------GAT----------HVIFCGLVPGAHYRVDIAS 458
```

```
Qy        437 VLGDIEGTPGFLQVHTPPVPVSDFRV-TVVSTTEIGLAW--SSHDAESFQMHITQEGAGN 493
Db        459 STGDISQS---ISGYTSPLPPQSLEVISRSSPSDLTIAWGPAPGQLEGYKVTWHQDGSQR 515

Qy        494 SRVEIT----TNQSIIIGGLFPGTKYCFEIVPKGPN-GTEGASRTVCNRTVPSAVFDIHV 548
Db        516 SPGDLVDLGPDTLSLTLKSLVPGSSYTVSAWAWAGNLGSDSQKIHSCTRPAPPTNLSLGF 575

Qy        549 VYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQG--LIPGTLYNIT- 605
Db        576 AH-QPAALKASWYHPPGGRD-AFHLRLYRL----------RPLTLESEKVLPREAQNFSW 623

Qy        606 --------ISPEVDHVWGDPNSTAQ----YTRPSNVSNIDVSTNT-TAATLSWQNFDDAS 652
Db        624 AQLTAGCEFQVQLSTLWGSERSSSANATGWTPPSAPTLVNVTSDAPTQLQVSWAHVPGGR 683

Qy        653 PTYSYCLLIEKAGNSSNATQVV-------TDIGITDATVTELIPGSSYTVEIFAQVGDGI 705
Db        684 SRYQVTLYQE---STRTATSIMGPKEDGTSFLGLT--------PGTKYKVEVISWAG--- 729

Qy        706 KSLEPGRKSFCTDPASMASFDCEVVPKE--------PALV-------------------- 737
Db        730 --------PLYTAAANVSAWTYPLIPNELLVSMQAGSAVVNLAWPSGPLGQGACHAQLSD 781

Qy        738 ---LKWTCP--------------PGANAGFELEVSSGAWNNATHLESCSSENG----TEY 776
Db        782 AGHLSWEQPLKLGQELFMLRDLTPGHTISMSVRCRAGPLQASTHLVVLSVEPGPVEDVLC 841

Qy        777 RTEVTYLNF-----------------------------------------------STSYNI 791
Db        842 HPEATYLALNWTMPAGDVDVCLVVVERLVPGGGTHFVFQVNTSGDALLLPNLMPTTSYRL 901

Qy        792 SITTVSCGKMAAPTRN---TCTTGITDPPPPD--GSPNITSVSHNSVKVKFSGFEASHGP 846
Db        902 SLTVL--GRNSRWSRAVSLVCSTSAEAWHPPELAEPPQVELGTGMGVTVMRGMFGKDDGQ 959

Qy        847 IKAYAVILTTG-EAGHPSADVLKYT-YDDFKKGASDTYVTYLIRT----EEKGRSQSLSE 900
Db        960 IQWYGIIATINMTLAQPSREAINYTWYDHYYRGC-ESFLALLFPNPFYPEPWAGPRS--- 1015

Qy        901 VLKYEIDVGNE--STTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFS 958
Db       1016 ---WTVPVGTEDCDNTQEICNGRLKSGFQYRFSVVAFSR----------LNTPETILAFS 1062

Qy        959 RYSD---AVSLPQDP-GVICGAVFGCIFGALVIVTVGGFIFWR-----KKRKDAKNNEVS 1009
Db       1063 AFSEPRASISLAIIPLTVMLGAVVGSI---VIVCAVLCLLRWRCLKGPRSEKDGFSKE-- 1117

Qy       1010 FSQIKP----KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAEN 1065
Db       1118 ---LMPYNLWRTHRPIPIHSFRQSYEAKSAHAHQTFFQEFEELKEVGKDQPRLEAEHPDN 1174

Qy       1066 RGKNRYNNVLPYDISRVKLSV---QTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKD 1122
Db       1175 IIKNRYPHVLPYDHSRVRLTQLPGEPHS--DYINANFIPGYSHTQEIIATQGPLKKTLED 1232

Qy       1123 FWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQ-DYGDITVAMTSEIVLPEWTIRD 1181
Db       1233 FWRLVWEQQVHVIIMLTVGMENGRVLCEHYWPANSTPVTHGHITIHLLAEEPEDEWTRRE 1292

Qy       1182 FTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSA 1241
Db       1293 FQLQHGTEQKQRRVKQLQFTTWPDHSVPEAPSSLLAFVELVQEQVQATQGKGPILVHCSA 1352

Qy       1242 GVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLD-IVR 1300
Db       1353 GVGRTGTFVALLRLLRQLEEEKVADVFNTVYILRLHRPLMIQTLSQYIFLHSCLLNKILE 1412

Qy       1301 SQKDS 1305
Db       1413 GPPDS 1417
```

RESULT 9
A49502
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type 4E, splice form A precursor - f
C;Species: Drosophila melanogaster
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: A49502
R;Oon, S.H.; Hong, A.; Yang, X.; Chia, W.
J. Biol. Chem. 268, 23964-23971, 1993
A;Title: Alternative splicing in a novel tyrosine phosphatase gene (DPTP4E) of Drosophila
A;Reference number: A49502; MUID:94043220; PMID:8226938
A;Accession: A49502
A;Molecule type: mRNA
A;Residues: 1-1767 <OON>
A;Cross-references: UNIPROT:Q9W4F5; UNIPROT:Q24495; GB:L20894
A;Note: authors translated the codon ATA for residue 1715 as Leu
C;Genetics:
A;Gene: ptp4E
A;Cross-references: FlyBase:FBgn0004368
C;Superfamily: protein-tyrosine-phosphatase, receptor type 4E; fibronectin type III repea
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recepto
F;1254-1270/Domain: transmembrane #status predicted <TMN>
F;1271-1767/Domain: intracellular #status predicted <INT>
F;1353-1573/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1525/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1531/Binding site: substrate phosphate (Arg) #status predicted

Query Match 13.6%; Score 954; DB 2; Length 1767;
Best Local Similarity 24.5%; Pred. No. 4.7e-45;
Matches 352; Conservative 179; Mismatches 448; Indels 458; Gaps 56;

```
Qy         67 SFHKQ-NGTGTP---------QVETNTSE-------------------DGESSGAND-- 94
Db        469 SYHEQTNASEVPAPFPVAAESQITTNLTEYTLDSLLAGRRYLIAVQALSKGVASNASDIT 528

Qy         95 -----------SLRTPEQG---------------------SNGTDGASQKTPSSTG---- 118
Db        529 RYTRPAAPLIQELRSIDQGLMLSWRSDVNSRQDRYEVHYQRNGTREERTMATNETSLTIH 588

Qy        119 ---PSPVFDIKAVSIS-----------------PTN----------VILTWKSNDTAAS 147
Db        589 YLHPGSGYEVKVHAISHGVRSEPHSYFQAVFPKPPQNLTLQTVHTNLVVLHWQAPE--GS 646

Qy        148 EY-KYVVKHKMENEKTITVVHQPWCNITGLRP----------ATSYVFSI-TPGIGNETW 195
Db        647 DFSEYVVRYRTD--------ASPWQRISGLHENEARIKDMHYGERYLVQVNTVSFGVES- 697

Qy        196 GDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASC-----------RVLLESI 243
Db        698 PHPLELNVTMPPQPVSNV-VPLVDSRNLTLEWPRPDGHVDFYTLKWWPTDEEDRVEFKNV 756

Qy        244 GSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSRAGS 303
Db        757 TQLEDLSSPS-VRIPIEDLSPGRQYRF-----------EVQASSNGIRSGTTHLSTRTM 803

Qy        304 PTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEP----GTRYNATVYSQAANGTEGQPQA 359
Db        804 PL--IQSDVFIA----NAGHEQGQDETITLSYTPTPADSTRFDIYRFSM------GDP-- 849

Qy        360 IEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419
Db        850 -------------------------TIKDKEKLANDT----------------ERKL 865

Qy        420 VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA 479
Db        866 SFSGLTPGKLYNVTVWTVSGGVASLPVQRVYRLHPLPISDLKAIQVAAREITLHWTAPAG 925

Qy        480 E----SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPN--GTEGASRT 533
Db        926 EYTDFELQYLSADEEAPQLLQNVTKNTEITLQGLRPYHNYTFTVVVRSGSIQGTDFADVS 985

Qy        534 VCNRTVPSAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITL 593
```

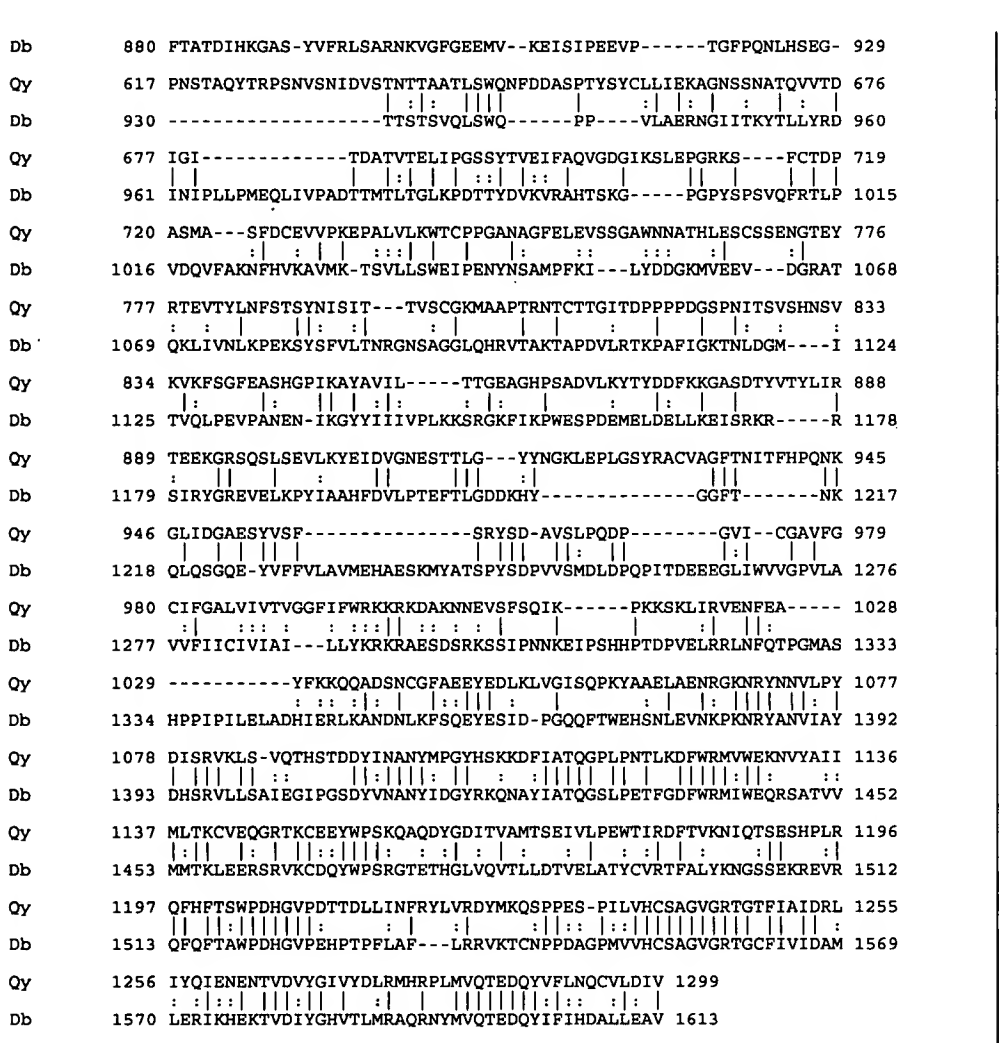
```
Db        986 VSTLMRSSA--PISASYQTLT-------APPGKVDYFQPSDVQ----------------- 1019

Qy        594 QGLIPGTLYNITISPEVDHVW----GDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFD 649

Db       1020 ----PG---------EVTFEWSLEPAEQHGPIDYFR----------------ITCQNAD 1049

Qy        650 DASPTYSYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLE 709

Db       1050 DAADVSSYEFPV-------NATQ---------GKIDGLVPGNHYIFRIQAKSALGY---- 1089

Qy        710 PGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCS 769

Db       1090 ----------------------------------------GAERE----------HIQ--- 1097

Qy        770 SENGTEYRTEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVS 829

Db       1098 ------------------------------------TMPILAPPVPEPSVTPLEVS 1117

Qy        830 HNSVKVKFS----GFEASHGPIKAYAVILT--TGEAGH----PS-ADVLKYT----YDDF 874

Db       1118 RTSSTIEISFRQGYFSNAHGMVRSYTIIIAEDVGKIASGLEMPSWQDVQAYTVWLPYQ-- 1175

Qy        875 KKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGN-ESTTLGYYNGKLEPLGSYRACVA 933

Db       1176 ---AIEPYNPFL--TSNGSRKSSL-EAEHLTIGTANCDKHQAGYCNGPLRAGTTYRIKIR 1229

Qy        934 GFTNITFHPQNKGLIDGAESYVSFSRYSDAVSLPQ-DPGVICGAVFGCIFGALVIVTVGG 992

Db       1230 AFTD-------------EDKFTDTVYSSPITTERSDTVIVAATVSAVLLVAMVLVVV-- 1273

Qy        993 FIFWRKKRKDAKNNEVSFSQIKPKKSKL----------IRVENFEAYFKKQQADSNCGFA 1042

Db       1274 ---YCQHRCQLIRRASKLARMQDELAALPEGYITPNRPVHVKDFSEHYRIMSADSDFRFS 1330

Qy       1043 EEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMP 1101

Db       1331 EEFEELKHVGRDQACSFANLPCNRPKNRFTNILPYDHSRFKLQPVDDDDGSDYINANYMP 1390

Qy       1102 GYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQD 1160

Db       1391 GHNSPREFIVTQGPFHSTREEFWRMCWESNSRAIVMLTRCFEKGREKCDQYWPVDRVAMF 1450

Qy       1161 YGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRY 1220

Db       1451 YGDIKVQLIIDTHYHDWSISEFMVS--RNCESRIMRHFHFTTWPDFGVPEPPLSLVRFVR 1508

Qy       1221 LVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPL 1280

Db       1509 AFRDVI--GTDMRPIIVHCSAGVGRSGTFIALDRILQHIHKSDYVDIFGIVFAMRKERVF 1566

Qy       1281 MVQTEDQYVFLNQCVLDIVRSQKDSKVDLI-------YQNTTAMTIYENLAPVTTFG 1330

Db       1567 MVQTEQQYVCIHQCLLAVLEGKEHLLADSLELHANDGYEVT---KIYLERQPQTKMG 1620


RESULT 10
B49502
protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type 4E, splice form B precursor - f
C;Species: Drosophila melanogaster
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: B49502
R;Oon, S.H.; Hong, A.; Yang, X.; Chia, W.
J. Biol. Chem. 268, 23964-23971, 1993
A;Title: Alternative splicing in a novel tyrosine phosphatase gene (DPTP4E) of Drosophil
A;Reference number: A49502; MUID:94043220; PMID:8226938
A;Accession: B49502
A;Molecule type: mRNA
A;Residues: 1-1615 <OON>
A;Cross-references: UNIPROT:Q9W4F5; GB:L20894
C;Genetics:
A;Gene: FlyBase:Ptp4E
A;Cross-references: FlyBase:FBgn0004368
A;Introns: 1605/3

C;Superfamily: protein-tyrosine-phosphatase, receptor type 4E; fibronectin type III repe
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept
F;1254-1270/Domain: transmembrane #status predicted <TMN>
F;1271-1615/Domain: intracellular #status predicted <INT>
F;1353-1573/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1525/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1531/Binding site: substrate phosphate (Arg) #status predicted

  Query Match             13.5%;  Score 949;  DB 2;  Length 1615;
  Best Local Similarity   24.5%;  Pred. No. 7.8e-45;
  Matches  345;  Conservative  177;  Mismatches  438;  Indels  448;  Gaps   54;

Qy         67 SFHKQ-NGTGTP---------QVETNTSE-------------------DGESSGAND-- 94

Db        469 SYHEQTNASEVPAPFPVAAESQITTNLTEYTLDSLLAGRRYLIAVQALSKGVASNASDIT 528

Qy         95 -----------SLRTPEQG---------------------SNGTDGASQKTPSSTG---- 118

Db        529 RYTRPAAPLIQELRSIDQGLMLSWRSDVNSRQDRYEVHYQRNGTREERTMATNETSLTIH 588

Qy        119 ---PSPVFDIKAVSIS-----------------PTN----------VILTWKSNDTAAS 147

Db        589 YLHPGSGYEVKVHAISHGVRSEPHSYFQAVFPKPPQNLTLQTVHTNLVVLHWQAPE--GS 646

Qy        148 EY-KYVVKHKMENEKTITVVHQPWCNITGLRP----------ATSYVFSI-TPGIGNETW 195

Db        647 DFSEYVVRYRTD--------ASPWQRISGLHENEARIKDMHYGERYLVQVNTVSFGVES- 697

Qy        196 GDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASC-----------RVLLESI 243

Db        698 PHPLELNVTMPPQPVSNV-VPLVDSRNLTLEWPRPDGHVDFYTLKWWPTDEEDRVEFKNV 756

Qy        244 GSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSRAGS 303

Db        757 TQLEDLSSPS-VRIPIEDLSPGRQYRF-----------EVQASSNGIRSGTTHLSTRTM 803

Qy        304 PTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEP----GTRYNATVYSQAANGTEGQPQA 359

Db        804 PL--IQSDVFIA----NAGHEQGQDETITLSYTPTPADSTRFDIYRFSM------GDP-- 849

Qy        360 IEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419

Db        850 ---------------------------TIKDKEKLANDT----------------ERKL 865

Qy        420 VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA 479

Db        866 SFSGLTPGKLYNVTVWTVSGGVASLPVQRVYRLHPLPISDLKAIQVAAREITLHWTAPAG 925

Qy        480 E----SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPN--GTEGASRT 533

Db        926 EYTDFELQYLSADEEAPQLLQNVTKNTEITLQGLRPYHNYTFTVVVRSGSIQGTDFADVS 985

Qy        534 VCNRTVPSAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITL 593

Db        986 VSTLMRSSA--PISASYQTLT-------APPGKVDYFQPSDVQ----------------- 1019

Qy        594 QGLIPGTLYNITISPEVDHVW----GDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFD 649

Db       1020 ----PG---------EVTFEWSLEPAEQHGPIDYFR----------------ITCQNAD 1049

Qy        650 DASPTYSYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLE 709

Db       1050 DAADVSSYEFPV-------NATQ---------GKIDGLVPGNHYIFRIQAKSALGY---- 1089

Qy        710 PGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCS 769

Db       1090 ----------------------------------------GAERE----------HIQ--- 1097

Qy        770 SENGTEYRTEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVS 829

Db       1098 ------------------------------------TMPILAPPVPEPSVTPLEVS 1117

Qy        830 HNSVKVKFS----GFEASHGPIKAYAVILT--TGEAGH----PS-ADVLKYT----YDDF 874
```

```
Db       1118 RTSSTIEISFRQGYFSNAHGMVRSYTIIIAEDVGKIASGLEMPSWQDVQAYTVWLPYQ-- 1175

Qy        875 KKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGN-ESTTLGYYNGKLEPLGSYRACVA 933

Db       1176 ---AIEPYNPFL--TSNGSRKSSL-EAEHLTIGTANCDKHQAGYCNGPLRAGTTYRIKIR 1229

Qy        934 GFTNITFHPQNKGLIDGAESYVSFSRYSDAVSLPQ-DPGVICGAVFGCIFGALVIVTVGG 992

Db       1230 AFTD-------------EDKFTDTVYSSPITTERSDTVIVAATVSAVLLVAMVLVVV-- 1273

Qy        993 FIFWRKKRKDAKNNEVSFSQIKPKKSKL----------IRVENFEAYFKKQQADSNCGFA 1042

Db       1274 ---YCQHRCQLIRRASKLARMQDELAALPEGYITPNRPVHVKDFSEHYRIMSADSDFRFS 1330

Qy       1043 EEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMP 1101

Db       1331 EEFEELKHVGRDQACSFANLPCNRPKNRFTNILPYDHSRFKLQPVDDDDGSDYINANYMP 1390

Qy       1102 GYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWP-SKQAQD 1160

Db       1391 GHNSPREFIVTQGPFHSTREEFWRMCWESNSRAIVMLTRCFEKGREKCDQYWPVDRVAMF 1450

Qy       1161 YGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRY 1220

Db       1451 YGDIKVQLIIDTHYHDWSISEFMVS--RNCESRIMRHFHFTTWPDFGVPEPPLSLVRFVR 1508

Qy       1221 LVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPL 1280

Db       1509 AFRDVI--GTDMRPIIVHCSAGVGRSGTFIALDRILQHIHKSDYVDIFGIVFAMRKERVF 1566

Qy       1281 MVQTEDQYVFLNQCVLDIVRSQKDSKVD 1308

Db       1567 MVQTEQQYVCIHQCLLAVLEGKEHLLAD 1594
```

RESULT 11
A56178
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type delta precursor - human
N;Alternate names: protein-tyrosine-phosphatase BPTP-2
C;Species: Homo sapiens (man)
C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56178; S12052; B44929
R;Pulido, R.; Krueger, N.X.; Serra-Pages, C.; Saito, H.; Streuli, M.
J. Biol. Chem. 270, 6722-6728, 1995
A;Title: Molecular characterization of the human transmembrane protein-tyrosine phosphat
ase delta isoforms.
A;Reference number: A56178; MUID:95204468; PMID:7896816
A;Accession: A56178
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1912 <PUL>
A;Cross-references: UNIPROT:P23468; GB:L38929; NID:g755652; PIDN:AAC41749.1; PID:g755653
R;Krueger, N.X.; Streuli, M.; Saito, H.
EMBO J. 9, 3241-3252, 1990
A;Title: Structural diversity and evolution of human receptor-like protein tyrosine phos
A;Reference number: S12049; MUID:91006018; PMID:2170109
A;Accession: S12052
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 390-1912 <KRU>
A;Cross-references: GB:X54133; NID:g35789; PIDN:CAA38068.1; PID:g35790
A;Note: the sequence from Fig. 5B is inconsistent with that from Fig. 5A in having 568-T
R;Adachi, M.; Sekiya, M.; Arimura, Y.; Takekawa, M.; Itoh, F.; Hinoda, Y.; Imai, K.; Yac
Cancer Res. 52, 737-740, 1992
A;Title: Protein-tyrosine phosphatase expression in pre-B cell NALM-6.
A;Reference number: A44929; MUID:92119637; PMID:1370651
A;Accession: B44929
A;Molecule type: mRNA
A;Residues: 1756-1804,'C',1806-1845 <ADA>
A;Cross-references: GB:S78086; NID:g243545; PIDN:AAB21147.1; PID:g243546
A;Experimental source: pre-B cell NALM-6
A;Note: sequence extracted from NCBI backbone (NCBIN:78086, NCBIP:78087)
A;Note: the authors did not report the entire codon for residue 90
C;Genetics:
A;Gene: GDB:PTPRD
A;Cross-references: GDB:131384; OMIM:601598
A;Map position: 9p24-9p24
C;Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology; i
ogy
C;Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane p
F;38-100/Domain: immunoglobulin homology <IMM1>
F;140-209/Domain: immunoglobulin homology <IMM2>
F;250-304/Domain: immunoglobulin homology <IMM3>
F;711-811/Domain: fibronectin type III repeat homology <3FR>
F;1293-1912/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F;1669-1892/Domain: protein-tyrosine-phosphatase homology <PTP2>
F;1553/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1559/Binding site: substrate phosphate (Arg) #status predicted
F;1844/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1850/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.5%; Score 879; DB 2; Length 1912;
Best Local Similarity 24.3%; Pred. No. 8.5e-41;
Matches 375; Conservative 209; Mismatches 560; Indels 400; Gaps 67;

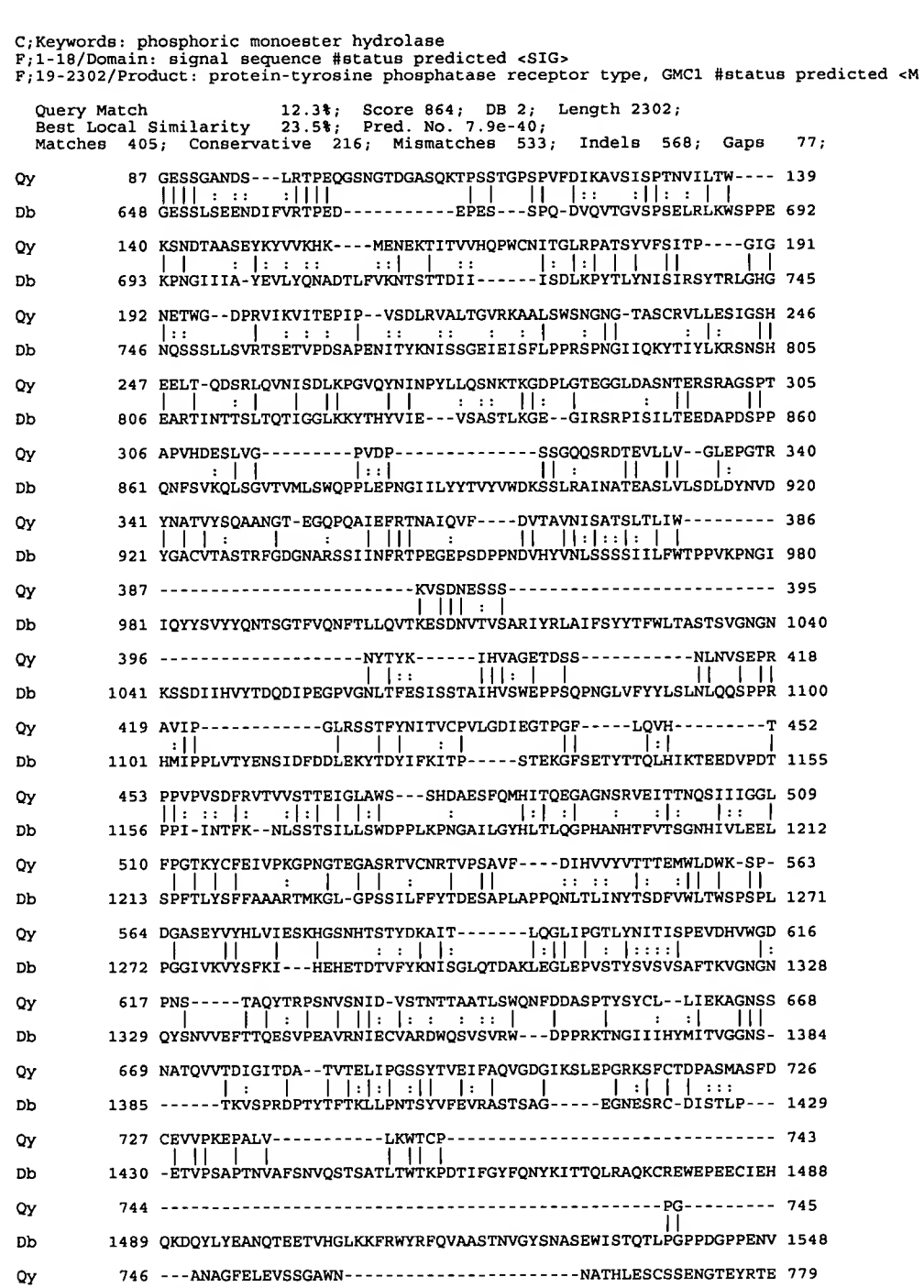
```
Qy         14 SPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQISSTAESFHKQNG 73

Db        212 SAGTRYSAPANLYVRELREVRRVPPRFS-IPPTNHEIMPGGSVNITCVA----------- 259

Qy         74 TGTPQ--------VETNTSEDGESSGAND-SLRTPEQGSNGT----------DGASQKTP 114

Db        260 VGSPMPYVKWMLGAEDLTPEDDMPIGRNVLELNDVRQSANYTCVAMSTLGVIEAIAQITV 319

Qy        115 SSTGPSPVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENE----KTITVVHQPW 170

Db        320 KAL-PKPPGTPVVTESTATSITLTWDSGNPEPVSY-YIIQHKPKNSEELYKEIDGVATTR 377

Qy        171 CNITGLRPATSYVFSITPGIGNETWGDPR--VIKVITEPIPVS---DLRVALTGVRKAAL 225

Db        378 YSVAGLSPYSDYEFRVV-AVNNIGRGPPSEPVLTQTSEQAPSSAPRDVQARMLSSTTILV 436

Qy        226 SW---SNGNGTASCRVLLESIGSHEELTQDSRLQVN--------------ISDLKPGVQY 268

Db        437 QWKEPEEPNGQIQ--------GYRVYYTMDPTQHVNNWMKHNVADSQITTIGNLVPQKTY 488

Qy        269 NINPYLLQSNKTKGDPLGTEGGLDASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDT 328

Db        489 SVKVLAFTS-------IG-DGPLSSDIQVITQTGVPGQPLNFKA-----EPES-----ET 530

Qy        329 EVLL----------------------------VGLEPGTRY-------NATVYSQAANGTE 354

Db        531 SILLSWTPPRSDTIANYELVYKDGEHGEEQRITIEPGTSYRLQGLKPNSLYYFRLA---A 587

Qy        355 GQPQAIEFRTNAIQV-----------FDVTAVNISATSLTLIWK----VSDNESSSNYTY 399

Db        588 RSPQGLGASTAEISARTMQSKPSAPPQDISCTSPSSTSILVSWQPPPVEKQNGIITEYSI 647

Qy        400 K-IHVAGETDSSNLNVSEP----RAVIPGLRSSTFYNITVCPVLGDIEGTPGFLQV---- 450

Db        648 KYTAVDGEDDKPHEILGIPSDTTKYLLEQLEKWTEYRITV-TAHTDVGPGPESLSVLIRT 706

Qy        451 --HTPPVPVSDFRVTVVSTTEIGLAWSS-------HDAESFQMHITQEGAGNSR------ 495

Db        707 NEDVPSGPPRKVEVEAVNSTSVKVSWRSPVPNKQHGQIRGYQVHYVRMENGEPKGQPMLK 766

Qy        496 ----------VEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFD 545

Db        767 DVMLADAQWEFDDTTEHDMIISGLQPETSYSLTVTAYTTKG-DGA------RSKPKLVST 819

Qy        546 IHVV------YVTTTEM---WLDWKSP-------------------------------- 563

Db        820 TGAVPGKPRLVINHTQMNTALIQWHPPVDTFGPLQGYRLKFGRKDMEPLTTLEFSEKEDH 879

Qy        564 -------DGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHVWGD 616
```

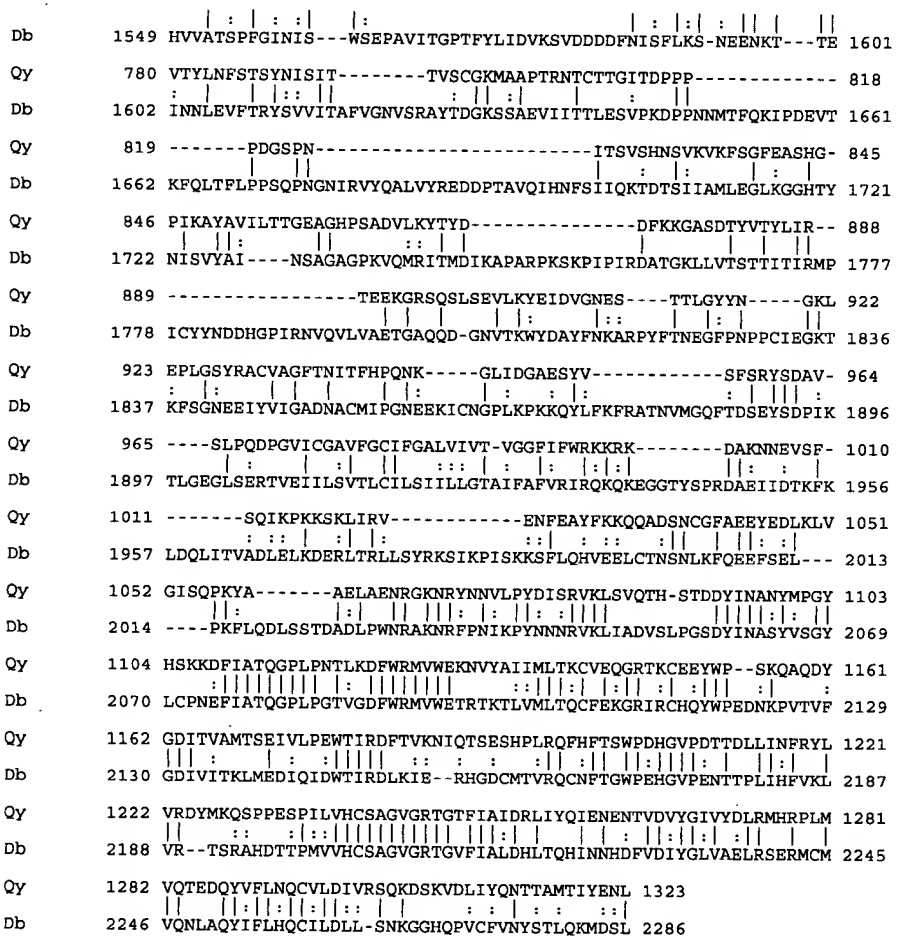


RESULT 12
T14328
protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type, GMC1 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14328
R;Wright, M.B.; Hugo, C.; Seifert, R.; Disteche, C.M.; Bowen-Pope, D.F.
J. Biol. Chem. 273, 23929-23937, 1998
A;Title: Proliferating and migrating mesangial cells responding to injury express a nove
A;Reference number: Z17986; MUID:98395110; PMID:9727007
A;Accession: T14328
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2302 <WRI>
A;Cross-references: UNIPROT:O88488; EMBL:AF063249; NID:g3300095; PID:g3300096; PIDN:AAC3
A;Experimental source: strain Wistar
C;Genetics:
A;Note: PTPRQ

C;Keywords: phosphoric monoester hydrolase
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-2302/Product: protein-tyrosine phosphatase receptor type, GMC1 #status predicted <M

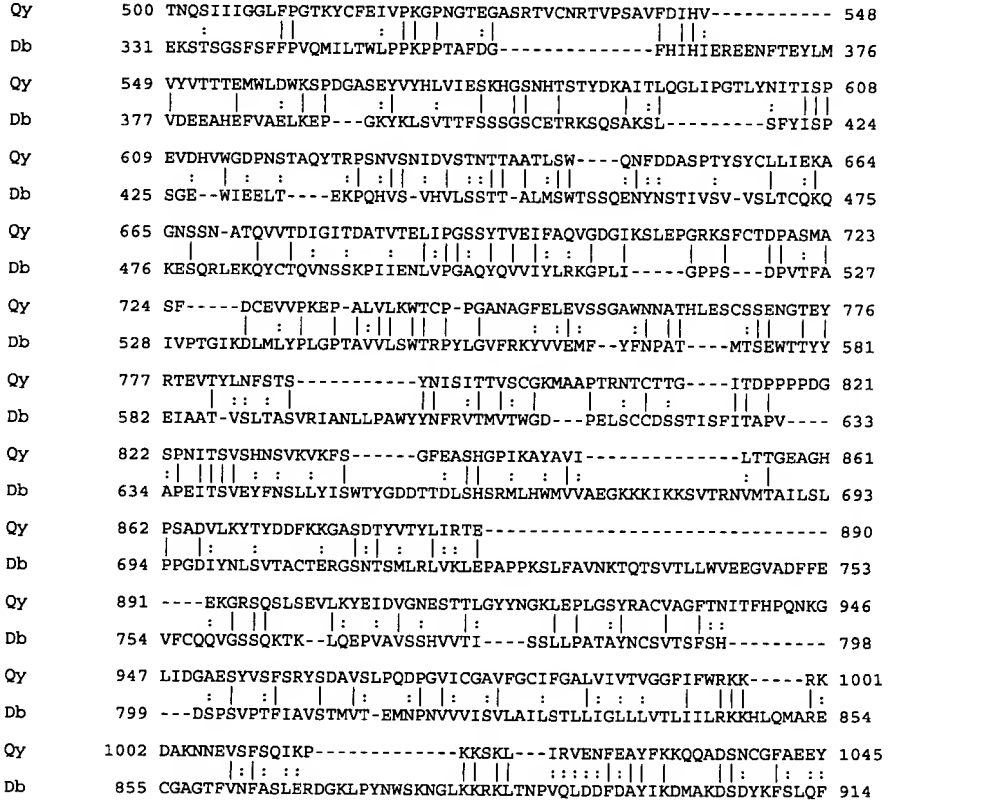
Query Match 12.3%; Score 864; DB 2; Length 2302;
Best Local Similarity 23.5%; Pred. No. 7.9e-40;
Matches 405; Conservative 216; Mismatches 533; Indels 568; Gaps 77;

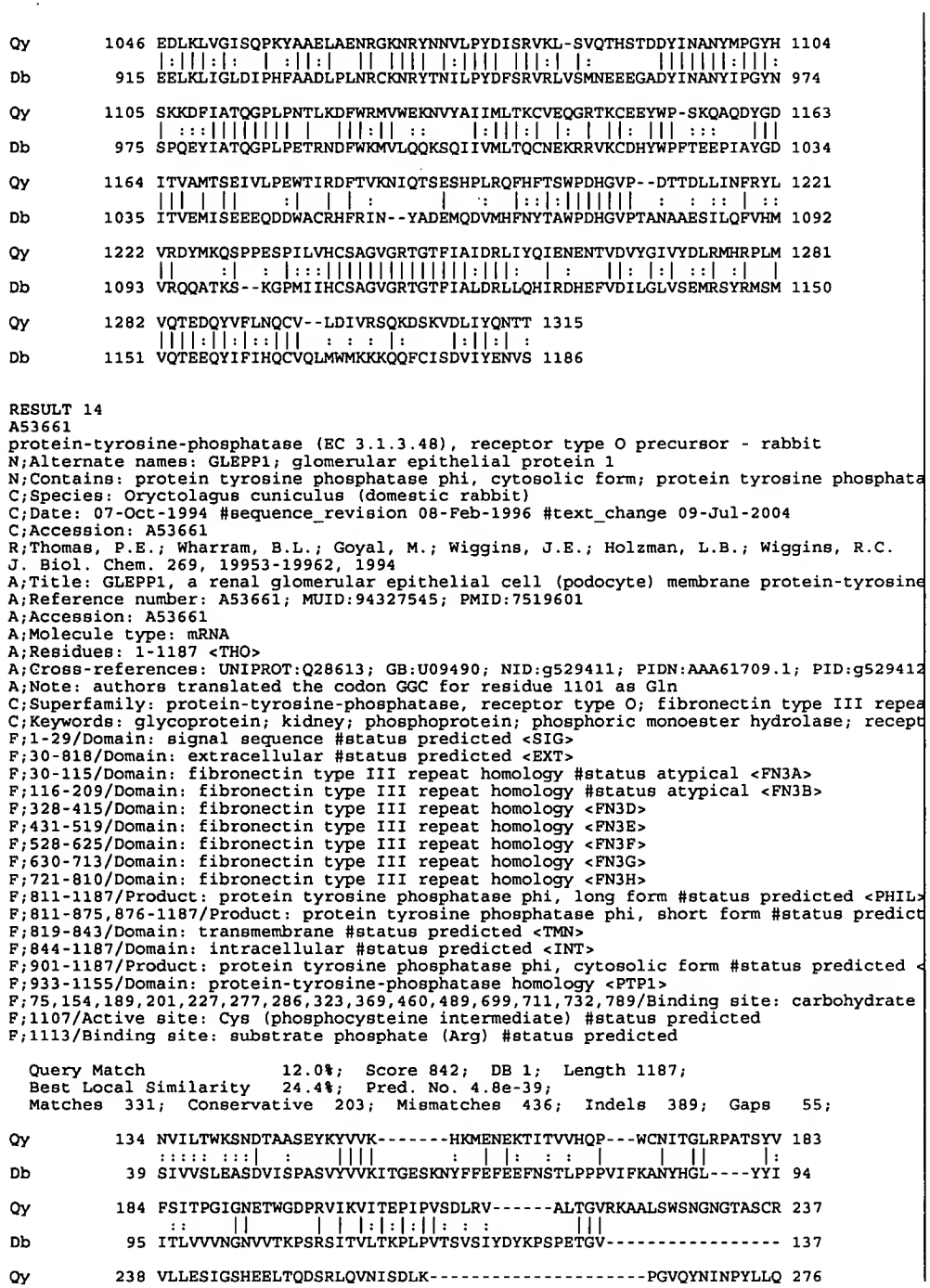




RESULT 13
A57064
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type O precursor - human
N;Alternate names: GLEPP1; glomerular epithelial protein 1
N;Contains: protein tyrosine phosphatase phi, cytosolic form; protein tyrosine phosphata
C;Species: Homo sapiens (man)
C;Date: 03-Oct-1995 #sequence_revision 08-Feb-1996 #text_change 22-Jun-1999
C;Accession: A57064
R;Wiggins, R.C.; Wiggins, J.E.; Goyal, M.; Wharram, B.L.; Thomas, P.E.
Genomics 27, 174-181, 1995
A;Title: Molecular cloning of cDNAs encoding human GLEPP1, a membrane protein tyrosine p
ne to human chromosome 12p12-p13.
A;Reference number: A57064; MUID:95394455; PMID:7665166
A;Accession: A57064
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1188 <WIG>
A;Cross-references: GB:U20489; NID:g885925; PIDN:AAA82892.1; PID:g885926
C;Genetics:
A;Gene: GDB:PTPRO
A;Cross-references: GDB:454477; OMIM:600579
A;Map position: 12p13.3-12p13.1
C;Superfamily: protein-tyrosine-phosphatase, receptor type O; fibronectin type III repeat
C;Keywords: glycoprotein; kidney; phosphoprotein; phosphoric monoester hydrolase; recepto
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-819/Domain: extracellular #status predicted <EXT>
F;30-109/Domain: fibronectin type III repeat homology #status atypical <FN3A>
F;116-202/Domain: fibronectin type III repeat homology #status atypical <FN3B>
F;329-409/Domain: fibronectin type III repeat homology <FN3C>
F;432-520/Domain: fibronectin type III repeat homology <FN3D>
F;529-619/Domain: fibronectin type III repeat homology <FN3F>
F;631-714/Domain: fibronectin type III repeat homology <FN3G>
F;722-804/Domain: fibronectin type III repeat homology <FN3H>
F;812-1188/Product: protein tyrosine phosphatase phi, long form #status predicted <PHIL>
F;812-876,877-1188/Product: protein tyrosine phosphatase phi, short form #status predicte
F;820-844/Domain: transmembrane #status predicted <TMN>
F;845-1188/Domain: intracellular #status predicted <INT>
F;902-1188/Product: protein tyrosine phosphatase phi, cytosolic form #status predicted <P
F;934-1156/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;75,154,189,201,227,278,287,323,324,370,461,490,700,712,733,790/Binding site: carbohydra
F;1108/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1114/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.1%; Score 854; DB 1; Length 1188;
Best Local Similarity 28.0%; Pred. No. 1e-39;
Matches 279; Conservative 158; Mismatches 347; Indels 212; Gaps 41;





RESULT 14
A53661
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type O precursor - rabbit
N;Alternate names: GLEPP1; glomerular epithelial protein 1
N;Contains: protein tyrosine phosphatase phi, cytosolic form; protein tyrosine phosphata
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 07-Oct-1994 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C;Accession: A53661
R;Thomas, P.E.; Wharram, B.L.; Goyal, M.; Wiggins, J.E.; Holzman, L.B.; Wiggins, R.C.
J. Biol. Chem. 269, 19953-19962, 1994
A;Title: GLEPP1, a renal glomerular epithelial cell (podocyte) membrane protein-tyrosine
A;Reference number: A53661; MUID:94327545; PMID:7519601
A;Accession: A53661
A;Molecule type: mRNA
A;Residues: 1-1187 <THO>
A;Cross-references: UNIPROT:Q28613; GB:U09490; NID:g529411; PIDN:AAA61709.1; PID:g529412
A;Note: authors translated the codon GGC for residue 1101 as Gln
C;Superfamily: protein-tyrosine-phosphatase, receptor type O; fibronectin type III repea
C;Keywords: glycoprotein; kidney; phosphoprotein; phosphoric monoester hydrolase; recept
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-818/Domain: extracellular #status predicted <EXT>
F;30-115/Domain: fibronectin type III repeat homology #status atypical <FN3A>
F;116-209/Domain: fibronectin type III repeat homology #status atypical <FN3B>
F;328-415/Domain: fibronectin type III repeat homology <FN3D>
F;431-519/Domain: fibronectin type III repeat homology <FN3E>
F;528-625/Domain: fibronectin type III repeat homology <FN3F>
F;630-713/Domain: fibronectin type III repeat homology <FN3G>
F;721-810/Domain: fibronectin type III repeat homology <FN3H>
F;811-1187/Product: protein tyrosine phosphatase phi, long form #status predicted <PHIL>
F;811-875,876-1187/Product: protein tyrosine phosphatase phi, short form #status predict
F;819-843/Domain: transmembrane #status predicted <TMN>
F;844-1187/Domain: intracellular #status predicted <INT>
F;901-1187/Product: protein tyrosine phosphatase phi, cytosolic form #status predicted <
F;933-1155/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;75,154,189,201,227,277,286,323,369,460,489,699,711,732,789/Binding site: carbohydrate
F;1107/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1113/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.0%; Score 842; DB 1; Length 1187;
Best Local Similarity 24.4%; Pred. No. 4.8e-39;
Matches 331; Conservative 203; Mismatches 436; Indels 389; Gaps 55;

```
Db       1031 YGDITVEMISEEEQDDWAHRHFRIN--YADEMQDVMHFNYTAWPDHGVPTANAAESILQF 1088

Qy       1219 RYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHR 1278

Db       1089 VHMVRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRDHEFVDILGLVSEMRSYR 1146

Qy       1279 PLMVQTEDQYVFLNQCV--LDIVRSQKDSKVDLIYQNTT 1315

Db       1147 MSMVQTEEQYIFIHQCVQLMWMKKKQQFCISDVIYENVS 1185


RESULT 15
S60613
protein-tyrosine-phosphatase (EC 3.1.3.48) U2 precursor - human
C;Species: Homo sapiens (man)
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S60613
R;Seimiya, H.; Sawabe, T.; Inazawa, J.; Tsuruo, T.
Oncogene 10, 1731-1738, 1995
A;Title: Cloning, expression and chromosomal localization of a novel gene for protein ty
A;Reference number: S60613; MUID:95273089; PMID:7753550
A;Accession: S60613
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1216 <SEI>
A;Cross-references: UNIPROT:Q16827; EMBL:Z48541; NID:g963058; PIDN:CAA88425.1; PID:g9630
C;Superfamily: protein-tyrosine-phosphatase, receptor type O; fibronectin type III repea
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatas
F;529-626/Domain: fibronectin type III repeat homology <3FR>
F;962-1184/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1136/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1142/Binding site: substrate phosphate (Arg) #status predicted

  Query Match             11.9%;  Score 839;  DB 2;  Length 1216;
  Best Local Similarity   24.2%;  Pred. No. 7.3e-39;
  Matches  333;  Conservative  200;  Mismatches  449;  Indels  394;  Gaps   57;

Qy       134 NVILTWKSNDTAASEYKYVVK-------HKMENEKTITVVHQP---WCNITGLRPATSYV 183

Db        39 NIVVSLEASDVISPASVYVVKITGESKNYFFEFEEFNSTLPPPVIFKASYHGL----YYI 94

Qy       184 FSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLLESI 243

Db        95 ITLVVVNGNVVTKPSRSITVLTKPLPVT------------------------------- 122

Qy       244 GSHEELTQDSRLQVNISDLKP----GVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERS 299

Db       123 ------------SVSIYDYKPSPETGVLFEIH-YPEKYNVFTRVNISYWGGKDFRTM--- 166

Qy       300 RAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQA 359

Db       167 --------LYKDFFKG-------------KTVFNHWLPGMCYSNITFQLVCEATFNKSTV 205

Qy       360 IEF----------RTNAI--QVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGET 407

Db       206 VEYSGVSHEPKQHRTAPYPPQNISVRIVNLNKNN----WE----EQSGNFPEESFMRSQD 257

Qy       408 DSSNLNVSEPRAVIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVS-------DF 460

Db       258 T---------------IGKEKLFHFT---------------EETPEIPSGNISSGWPDF 286

Qy       461 RVTVVSTTEIGLAWSSHDA-----ESFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKY 515

Db       287 NSSDYETTSQPYWWDSASAAPESEDEFFSVLPMEYENNSTLSETEKSTSGSFSFFPVQMI 346

Qy       516 CFEIVPKGPNGTEGASRTVCNRTVPSAVPDIHV-----------VYVTTTEMWLDWKSPD 564

Db       347 LTWLPPKPPTAFDG-------------FHIHIERBENFTEYLMVDEEAHEFVAELKEP- 391

Qy       565 GASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHVWGDPNSTAQYT 624

Db       392 --GKYKLSVTTFSSSGSCETRKSQSAKSL---------SFYISPSGE--WIEELT----E 434

Qy       625 RPSNVSNIDVSTNTTAATLSW----QNFDDASPTYSYCLLIEKAGNSSN-ATQVVTDIGI 679

Db       435 KPQHVS-FHVLSSTT-ALMSWTSSQENYNSTIVSV-VSLTCQKQKESQRLEKQYCTQVNS 491

Qy       680 TDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPASMASF-----DCEVVPKEP 734

Db       492 SKPIIENLVPGAQYQVVIYLRKGPLI-----GPPS---DPVTFAIVPTGIKDLMLYPLGP 543

Qy       735 -ALVLKWTCP-PGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEVTYLNFSTS---- 788

Db       544 TAVVLSWTRPYLGVFRKYVVEMF--YFNPAT----MTSEWTTYYEIAAT-VSLTASVRIA 596

Qy       789 -------YNISITTVSCGKMAAPTRNTCTTG----ITDPPPPDGSPNITSVSHNSVKVKF 837

Db       597 NLLPAWYYNFRVTMVTWGD---PELSCCDSSTISFITAPV----APEITSVEYFNSLLYI 649

Qy       838 S------GFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGAS------------ 879

Db       650 SWTYGDDTTDLSHSRMLHWMVV-TEGK------------KKIKKSVTRNVMTAILSLPP 695

Qy       880 -DTYVTYLIRTEEKGRSQSLSEVLKYE---------IDVGNESTTLGYYN---------- 919

Db       696 GDTYNLSVTTCTERGSNTSMLRLVKLEPAPPKSLFAVNKTQTSVTLLWVEEGVADFFKVF 755

Qy       920 ------------------------GKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESY 954

Db       756 FQHVGSSQKTKLQEPVAVSPHVVTISSLLPATAYSCSVTSFSH-----------DSPSV 803

Qy       955 VSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKK-------------- 999

Db       804 PTFIAVSTMVT-EMNPNVVVISVLAILSTLLIGLLLVTLIILRKKHLQMARECGAGTFAN 862

Qy      1000 ----RKDAK------NNEVSFSQIKP--------------------KKSKL---IRVEN 1025

Db       863 CASLERDGKLPYNCRRSIFAFLTLLPSCLWTDYPLAFYINPWSKNGLKKRKLTNPVQLDD 922

Qy      1026 FEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKL- 1084

Db       923 FDAYIKDMAKDSDYKFSLQFEELKLIGLDIPHFAADLPLNRCKNRYTNILPYDFSRVRLV 982

Qy      1085 SVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQ 1144

Db       983 SMNEEEGADYINANYIPGYNSPQEYIATQGPLPETRNDFWKMVLQQKSQIIVMLTQCNEK 1042

Qy      1145 GRTKCEEYWP-SKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSW 1203

Db      1043 RRVKCDHYWPFTEEPIAYGDITVEMISEEEQDDWACRHFRIN--YADEMQDVMHFNYTAW 1100

Qy      1204 PDHGVP--DTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIEN 1261

Db      1101 PDHGVPTANAAESILQFVHMVRQQATKS--KGPMIIHCSAGVGRTGTFIALDRLLQHIRD 1158

Qy      1262 ENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCV--LDIVRSQKDSKVDLIYQNTT 1315

Db      1159 HEFVDILGLVSEMRSYRMSMVQTEEQYIFIHQCVQLMWMKKKQQFCISDVIYENVS 1214


Search completed: May 10, 2005, 19:05:40
Job time : 69 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2005, 18:46:26 ; Search time 203 Seconds
(without alignments)
3372.659 Million cell updates/sec

Title: US-09-516-728A-4
Perfect score: 7030
Sequence: 1 MKPAAREARLPPRSPGLRWA.....TIYENLAPVTTFGKTNGYIA 1337

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7023	99.9	1337	1	PTPJ_HUMAN	Q12913 homo sapien
2	4510.5	64.2	1238	2	Q8K3Q2	Q8k3q2 mus musculu
3	4508.5	64.1	1238	1	PTPJ_MOUSE	Q64455 mus musculu
4	4499.5	64.0	1216	2	Q62884	Q62884 rattus norv
5	4483.5	63.8	1238	2	Q8CIW9	Q8ciw9 mus musculu
6	2784	39.6	539	2	Q6P4H4	Q6p4h4 homo sapien
7	2752.5	39.2	1406	2	Q9W6V5	Q9w6v5 gallus gall
8	2184	31.1	425	2	Q9NPR5	Q9npr5 homo sapien
9	1847	26.3	361	2	Q61373	Q61373 mus musculu
10	1259.5	17.9	1998	2	Q8CIW2	Q8ciw2 mus musculu
11	1242	17.7	1997	1	PTPB_HUMAN	P23467 homo sapien
12	1239	17.6	1956	2	Q6MZF6	Q6mzf6 homo sapien
13	1219.5	17.3	1102	2	Q80VN7	Q80vn7 mus musculu
14	1176.5	16.7	1115	2	Q9HD43	Q9hd43 homo sapien
15	1130	16.1	1118	2	Q15426	Q15426 homo sapien
16	1046.5	14.9	1631	2	Q86NN9	Q86nn9 drosophila
17	1043.5	14.8	1556	2	Q8IR87	Q8ir87 drosophila
18	1043.5	14.8	1631	1	PTP1_DROME	P35992 drosophila
19	1043.5	14.8	1962	2	Q9VYW1	Q9vyw1 drosophila
20	993.5	14.1	579	2	Q9JJ07	Q9jj07 mus musculu
21	991.5	14.1	1705	2	Q9ERK5	Q9erk5 mus musculu
22	990.5	14.1	1705	1	PTPV_MOUSE	P70289 mus musculu
23	969	13.8	1767	2	Q9W4F5	Q9w4f5 drosophila
24	964	13.7	1607	2	Q8IRS0	Q8irs0 drosophila
25	956	13.6	1711	1	PTPV_RAT	Q64612 rattus norv
26	955	13.6	1767	2	Q24495	Q24495 drosophila
27	922	13.1	1064	2	Q8SXB2	Q8sxb2 drosophila
28	906	12.9	598	2	Q8BIW7	Q8biw7 m mus muscu
29	879	12.5	1912	1	PTPD_HUMAN	P23468 homo sapien
30	864	12.3	2302	2	O88488	O88488 rattus norv
31	853.5	12.1	1198	2	Q9JLU0	Q9jlu0 mus musculu
32	852	12.1	1267	2	Q98945	Q98945 gallus gall
33	849.5	12.1	1216	2	Q62797	Q62797 rattus norv
34	846	12.0	1896	2	Q9IAJ1	Q9iaj1 xenopus lae
35	842	12.0	1187	2	Q28613	Q28613 oryctolagus
36	841	12.0	1226	2	Q7TSY7	Q7tsy7 mus musculu
37	839	11.9	1216	1	PTPO_HUMAN	Q16827 homo sapien
38	834.5	11.9	1226	2	Q9ERM5	Q9erm5 mus musculu
39	826	11.7	1894	2	Q64487	Q64487 mus musculu
40	822.5	11.7	1788	2	Q9IAJ0	Q9iaj0 xenopus lae
41	817	11.6	1367	2	Q20120	Q20120 caenorhabdi
42	808.5	11.5	1887	2	Q9QW67	Q9qw67 rattus sp.
43	807	11.5	1898	2	Q64604	Q64604 r protein-t
44	806.5	11.5	1898	2	Q9EQ17	Q9eq17 mus musculu
45	804	11.4	1948	1	PTNS_HUMAN	Q13332 homo sapien

ALIGNMENTS

RESULT 1
PTPJ_HUMAN
ID PTPJ_HUMAN STANDARD; PRT; 1337 AA.
AC Q12913; Q15255; Q8NHM2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Protein-tyrosine phosphatase eta precursor (EC 3.1.3.48) (R-PTP-eta)
DE (HPTP eta) (Protein-tyrosine phosphatase receptor type J) (Density
DE enhanced phosphatase-1) (DEP-1) (CD148 antigen).
GN Name=PTPRJ; Synonyms=DEP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95024024; PubMed=7937872;
RA Oestman A., Yang Q., Tonks N.K.;
RT "Expression of DEP-1, a receptor-like protein-tyrosine-phosphatase, is
RT enhanced with increasing cell density.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9680-9684(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95086212; PubMed=7994032;
RA Honda H., Inazawa J., Nishida J., Yazaki Y., Hirai H.;
RT "Molecular cloning, characterization, and chromosomal localization of
RT a novel protein-tyrosine phosphatase, HPTP eta.";
RL Blood 84:4186-4194(1994).
RN [3]
RP SEQUENCE OF 33-1337 FROM N.A., AND VARIANTS COLON CANCER CYS-214 AND
RP PRO-276.
RC TISSUE=Colon;
RX MEDLINE=22084388; PubMed=12089527; DOI=10.1038/ng903;
RA Ruivenkamp C.A.L., van Wezel T., Zanon C., Stassen A.P.M., Vlcek C.,
RA Csikos T., Klous A.M., Tripodis N., Perrakis A., Boerrigter L.,
RA Groot P.C., Lindeman J., Mooi W.J., Meijjer G.A., Scholten G.,
RA Dauwerse H., Paces V., van Zandwijk N., van Ommen G.J.B., Demant P.;
RT "Ptprj is a candidate for the mouse colon-cancer susceptibility locus
RT Scc1 and is frequently deleted in human cancers.";
RL Nat. Genet. 31:295-300(2002).
CC -!- FUNCTION: May contribute to the mechanism of contact inhibition of
CC cell growth.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- PTM: N- and O-glycosylated.
CC -!- DISEASE: Defects in PTPRJ are found in cancers of colon, lung, and
CC breast.
CC -!- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Receptor class 3 subfamily.
CC -!- SIMILARITY: Contains 9 fibronectin type III domains.
CC -!- SIMILARITY: Contains 1 protein-tyrosine phosphatase domain.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD148 entry;

```
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd148.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10886; AAB36687.1; -.
DR   EMBL; D37781; BAA07035.1; -.
DR   EMBL; AF387844; AAM69432.1; -.
DR   EMBL; AF387823; AAM69432.1; JOINED.
DR   EMBL; AF387824; AAM69432.1; JOINED.
DR   EMBL; AF387825; AAM69432.1; JOINED.
DR   EMBL; AF387826; AAM69432.1; JOINED.
DR   EMBL; AF387827; AAM69432.1; JOINED.
DR   EMBL; AF387828; AAM69432.1; JOINED.
DR   EMBL; AF387829; AAM69432.1; JOINED.
DR   EMBL; AF387830; AAM69432.1; JOINED.
DR   EMBL; AF387831; AAM69432.1; JOINED.
DR   EMBL; AF387832; AAM69432.1; JOINED.
DR   EMBL; AF387833; AAM69432.1; JOINED.
DR   EMBL; AF387834; AAM69432.1; JOINED.
DR   EMBL; AF387835; AAM69432.1; JOINED.
DR   EMBL; AF387836; AAM69432.1; JOINED.
DR   EMBL; AF387837; AAM69432.1; JOINED.
DR   EMBL; AF387838; AAM69432.1; JOINED.
DR   EMBL; AF387839; AAM69432.1; JOINED.
DR   EMBL; AF387840; AAM69432.1; JOINED.
DR   EMBL; AF387841; AAM69432.1; JOINED.
DR   EMBL; AF387842; AAM69432.1; JOINED.
DR   EMBL; AF387843; AAM69432.1; JOINED.
DR   PIR; I38670; I38670.
DR   HSSP; P18052; 1YFO.
DR   Genew; HGNC:9673; PTPRJ.
DR   MIM; 600925; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005001; F:transmembrane receptor protein tyrosine pho. . .; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR   GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; TAS.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 6.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 8.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 9.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Disease mutation; Glycoprotein; Hydrolase; Protein phosphatase;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     35       Potential.
FT   CHAIN        36   1337       Protein-tyrosine phosphatase eta.
FT   DOMAIN       36    975       Extracellular (Potential).
FT   TRANSMEM    976    996       Potential.
FT   DOMAIN      997   1337       Cytoplasmic (Potential).
FT   DOMAIN      119    205       Fibronectin type-III 1.
FT   DOMAIN      207    291       Fibronectin type-III 2.
FT   DOMAIN      271    364       Fibronectin type-III 3.
FT   DOMAIN      365    452       Fibronectin type-III 4.
FT   DOMAIN      453    538       Fibronectin type-III 5.
FT   DOMAIN      540    620       Fibronectin type-III 6.
FT   DOMAIN      622    717       Fibronectin type-III 7.
FT   DOMAIN      720    811       Fibronectin type-III 8.
FT   DOMAIN      816    902       Fibronectin type-III 9.
FT   DOMAIN     1065   1337       Protein-tyrosine phosphatase.
FT   ACT_SITE   1239   1239       Phosphocysteine intermediate (By
FT                                similarity).
FT   CARBOHYD     72     72       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     82     82       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     93     93       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    104    104       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    142    142       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    172    172       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    192    192       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    231    231       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    258    258       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    278    278       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    342    342       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    351    351       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    376    376       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    391    391       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    396    396       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    413    413       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    431    431       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    501    501       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    525    525       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    536    536       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    582    582       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    603    603       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    618    618       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    628    628       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    637    637       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    666    666       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    669    669       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    761    761       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    772    772       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    784    784       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    790    790       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    824    824       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    910    910       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    937    937       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT     214    214       R -> C (in colon cancer; somatic
FT                                mutation).
FT                                /FTId=VAR_015905.
FT   VARIANT     276    276       Q -> P (in colon cancer; somatic
FT                                mutation; dbSNP:1566734).
FT                                /FTId=VAR_015906.
FT   CONFLICT    261    261       G -> D (in Ref. 1).
FT   CONFLICT    918    929       YNGKLEPLGSYR -> LQWEAGTSGLLP (in Ref. 2).
SQ   SEQUENCE   1337 AA;  145926 MW;  E6752D521C4B6AFE CRC64;

  Query Match             99.9%;  Score 7023;  DB 1;  Length 1337;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1336;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISGLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
```

```
Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCTDPA 720

Qy        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEV 780

Qy        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGF 840

Qy        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSE 900

Qy        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRY 960

Qy        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL 1020

Qy       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 IRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDIS 1080

Qy       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTK 1140

Qy       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHF 1200

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Qy       1321 ENLAPVTTFGKTNGYIA 1337
              |||||||||||||||||
Db       1321 ENLAPVTTFGKTNGYIA 1337


RESULT 2
Q8K3Q2
ID   Q8K3Q2      PRELIMINARY;      PRT;  1238 AA.
AC   Q8K3Q2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   PTPRJ.
GN   Name=Ptprj;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=STS;
RA   van Wezel T., Ruivenkamp C.A.L., Zanon C., Stassen A.P.M., Vlcek C.,
RA   Csikos T., Tripodis N., Groot P.C., Dauwerse H., van Ommen G.J.B.,
RA   Demant P.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY039232; AAK98640.1; -.
DR   HSSP; P18052; 1YFO.
DR   MGD; MGI:104574; Ptprj.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   GO; GO:0007507; P:heart development; IMP.
DR   GO; GO:0001570; P:vasculogenesis; IMP.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 6.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 7.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 5.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase.
SQ   SEQUENCE   1238 AA;  136772 MW;  0D02F5EF8E23C0B2 CRC64;

  Query Match           64.2%;  Score 4510.5;  DB 2;  Length 1238;
  Best Local Similarity 67.2%;  Pred. No. 3.6e-247;
  Matches  901;  Conservative  108;  Mismatches  226;  Indels  105;  Gaps   12;

Qy          1 MKPAARERARLPPRSPGLRWALPLLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||| | |||||||||||  |||||| ||:|||| |:||
Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLRQGQVLCAGAAPNPI---------------- 43

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Db         44 ------------------------------------------------------------ 43

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
                |||:|| :|||:|:|||| ||: ||| :  :::|||:  |  | :|  |||||| | |
Db         44 --FDIEAV-VSPTSVLLTWKHNDSGASECR--IENKMESNLTFPVKNQTSCNITGLSPGT 98

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              || |||     |||        : ||| ||||| |   || :| |:||| ||||| |:|:
Db         99 SYTFSIISVTTNETLNK----TITTEPWPVSDLHVTSVGVTQARLTWSNANGTASYRMLI 154

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLL-QSNKTKGDPLGTEGGLDASNTERS 299
              | : :|          |||| ||||     | :   :||:|: |    |   ||: |:|
Db        155 EELTTHS--------SVNISGLKPGTN---NTFAFPESNETQADFAVAEEVPDANGTKRI 203

Qy        300 RAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQA 359
                : : | |  ||   ||| |||    ||:||  |:| |:|||:||||||||||||:
Db        204 PVTNLSQP-HKNSLAS-VDPPSGQDPSLTEILLTDLKPDTQYNATIYSQAANGTEGQPRN 261

Qy        360 IEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419
              |:||: || || |:||||:|:|| || : : | ::  ||||||| | | |  |::  |
Db        262 KVFKTNSTQVSDVRAMNISASSMTLTWKSNYDGSRTSIVYKIHVAGGTHSVNQTVNKTEA 321

Qy        420 VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA 479
```

```
Db        322 IILGLSSSTLYNITVHPFLGQTEGTPGFLQVYTSPDQVSDFRVTNVSTRAIGLAWRSNDS 381

Qy        480 ESFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTV 539

Db        382 KSFEIFIKQDGGEKHRNASTGNQSYMVEDLKPGTSYHFEIIPRGPDGTEGLSSTVNGSTD 441

Qy        540 PSAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGS-NHTSTYDKAITLQGLIP 598

Db        442 PSAVTDIRVVNISTTEMQLEWQNTDDASGYTYHLVLESKSGSIIRTNSSQKWITVGSLTP 501

Qy        599 GTLYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYC 658

Db        502 GTLYNVTIFPEVDQIQGISNSITQYTRPSSVSHIEVNTTTTTAAIRWKNEDTASASYAYS 561

Qy        659 LLIEKAGNSSNATQVVT-DIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCT 717

Db        562 VLILKTGDGSNVTSNFTKDPSI---LIPELIPGVSYTVKILTQVGDGTTSLVPGWNLFCT 618

Qy        718 DPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYR 777

Db        619 EPESVTSFHCEVVPKEPALVLKWACPFGMYTGFELGVRSDSWDNMTRLENCTSDDDTECR 678

Qy        778 TEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKF 837

Db        679 TEVAYLNFSTSYNISIATLSCGKMALPAQNICTTGITDPPTPDGSPNITSVSHNSVKVKF 738

Qy        838 SGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQS 897

Db        739 SGFEASHGPIKAYAVILTTGEAAQPSADVLKYTYEDFKRGASDTYVTYLIRIEEKGQSQG 798

Qy        898 LSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSF 957

Db        799 LSEVLNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYVSF 858

Qy        958 SRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKK 1017

Db        859 SPYSEAVFLPQDPGVICGAVFGCIFGALAITAVGGFIFWRKKRTDAKNNEVSFSQIKPKK 918

Qy       1018 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPY 1077

Db        919 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYTAEIAENRGKNRYNNVLPY 978

Qy       1078 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIM 1137

Db        979 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIVM 1038

Qy       1138 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQ 1197

Db       1039 LTKCVEQGRTKCEEYWPSKQAQEYGDITVAMTSEVVLPEWTIRDFVVKNMQNSESHPLRQ 1098

Qy       1198 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIY 1257

Db       1099 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRLIY 1158

Qy       1258 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAM 1317

Db       1159 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTTAM 1218

Qy       1318 TIYENLAPVTTFGKTNGYIA 1337

Db       1219 TIYENLEPVSMFGKTNGYIA 1238


RESULT 3
PTPJ_MOUSE
ID   PTPJ_MOUSE     STANDARD;      PRT;  1238 AA.
AC   Q64455;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Protein-tyrosine phosphatase eta precursor (EC 3.1.3.48) (R-PTP-eta)
DE   (HPTP beta-like tyrosine phosphatase) (Protein-tyrosine phosphatase
DE   receptor type J) (Susceptibility to colon cancer-1).
GN   Name=Ptprj; Synonyms=Byp, Scc1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MRL-LPR/LPR;
RX   MEDLINE=96140699; PubMed=8549806; DOI=10.1016/0014-5793(95)01415-2;
RA   Kuramochi S., Matsuda S., Matsuda Y., Saitoh T., Ohsugi M.,
RA   Yamamoto T.;
RT   "Molecular cloning and characterization of Byp, a murine receptor-type
RT   tyrosine phosphatase similar to human DEP-1.";
RL   FEBS Lett. 378:7-14(1996).
CC   -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC       tyrosine + phosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in every tissue examined.
CC   -!- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC       Receptor class 3 subfamily.
CC   -!- SIMILARITY: Contains 8 fibronectin type III domains.
CC   -!- SIMILARITY: Contains 1 protein-tyrosine phosphatase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D45212; BAA08146.1; -.
DR   PIR; S68700; S68700.
DR   HSSP; P18052; 1YFO.
DR   MGD; MGI:104574; Ptprj.
DR   GO; GO:0007507; P:heart development; IMP.
DR   GO; GO:0001570; P:vasculogenesis; IMP.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 6.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 7.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 7.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Glycoprotein; Hydrolase; Protein phosphatase; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL        1     28       Potential.
FT   CHAIN        29   1238       Protein-tyrosine phosphatase eta.
FT   DOMAIN       29    876       Extracellular (Potential).
FT   TRANSMEM    877    897       Potential.
FT   DOMAIN      898   1238       Cytoplasmic (Potential).
FT   DOMAIN       40    130       Fibronectin type-III 1.
FT   DOMAIN      170    266       Fibronectin type-III 2.
FT   DOMAIN      267    354       Fibronectin type-III 3.
FT   DOMAIN      356    440       Fibronectin type-III 4.
FT   DOMAIN      442    527       Fibronectin type-III 5.
FT   DOMAIN      529    618       Fibronectin type-III 6.
FT   DOMAIN      620    712       Fibronectin type-III 7.
FT   DOMAIN      717    803       Fibronectin type-III 8.
FT   DOMAIN      966   1238       Protein-tyrosine phosphatase.
FT   ACT_SITE   1140   1140       Phosphocysteine intermediate (By
FT                                similarity).
FT   CARBOHYD     62     62       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     78     78       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     85     85       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     90     90       N-linked (GlcNAc. . .) (Potential).
```

```
FT   CARBOHYD    110    110       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    114    114       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    145    145       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    164    164       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    173    173       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    182    182       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    198    198       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    207    207       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    244    244       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    253    253       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    267    267       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    278    278       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    313    313       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    317    317       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    333    333       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    366    366       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    379    379       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    398    398       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    403    403       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    437    437       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    452    452       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    488    488       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    506    506       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    538    538       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    572    572       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    576    576       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    662    662       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    668    668       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    685    685       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    691    691       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    725    725       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    811    811       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    838    838       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1238 AA;  136782 MW;  939479EDC8016835 CRC64;

  Query Match             64.1%;  Score 4508.5;  DB 1;  Length 1238;
  Best Local Similarity   67.2%;  Pred. No. 4.7e-247;
  Matches  901;  Conservative  108;  Mismatches  226;  Indels  105;  Gaps   12;

Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              |||||| | |||||||||||  |||||| ||:|||| |:||
Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLLRQGQVLCAGAAPNPI---------------- 43

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Db         44 ------------------------------------------------------------ 43

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
               |||:|| :|||:|:|||| ||: ||| :  :::|||:  |  | :|  |||||| | |
Db         44 --FDIEAV-VSPTSVLLTWKHNDSGASECR--IENKMESNLTFPVKNQTSCNITGLSPGT 98

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
               || |||     |||        : ||| ||||| |   || :| |:||| ||||| |:|:
Db         99 SYTFSIISVTTNETLNK----TITTEPWPVSDLHVTSVGVTQARLTWSNANGTASYRMLI 154

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLL-QSNKTKGDPLGTEGGLDASNTERS 299
               | : :|          |||| ||||     | :   :||:|: |    |   ||: |:|
Db        155 EELTTHS--------SVNISGLKPGTN---NTFAFPESNETQADFAVAEEVPDANGTKRI 203

Qy        300 RAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQA 359
                 :  : :|  |||  ||| |||    ||:||  |:| |:||||:||||||||||:
Db        204 PV-TNLSQLHKNSLVS-VDPPSGQDPSLTEILLTDLKPDTQYNATIYSQAANGTEGQPRN 261

Qy        360 IEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419
               |:||: || || |:||||:|:|| || : : | ::  |||||| | | |  |::  |
Db        262 KVFKTNSTQVSDVRAMNISASSMTLTWKSNYDGSRTSIVYKIHVAGGTHSVNQTVNKTEA 321

Qy        420 VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA 479
               :| || ||| ||||| | ||  |||||||||:| |  |||||| |||  ||||| |:|:
Db        322 IILGLSSSTLYNITVHPFLGQTEGTPGFLQVYTSPDQVSDFRVTNVSTRAIGLAWRSNDS 381

Qy        480 ESFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTV 539
               :||:: | |:|    |   | ||| ::  | ||| | |||:|:||:|||| | ||   |

Db        382 KSFEIFIKQDGGEKHRNASTGNQSYMVEDLKPGTSYHFEIIPRGPDGTEGLSSTVNGSTD 441

Qy        540 PSAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGS-NHTSTYDKAITLQGLIP 598
              |||| || || ::|||| |:|:: | || | ||||:||| ||    |::  | ||:  | |
Db        442 PSAVTDIRVVNISTTEMQLEWQNTDDASGYTYHLVLESKSGSSIIRTNSSQKWITVGSLTP 501

Qy        599 GTLYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYC 658
              |||||:|| |||| : |  ||  ||||||:||:|:|:| || | : |:| | || :|:|
Db        502 GTLYNVTIFPPEVDQIQGISNSITQYTRPSSVSHIEVNTTTTTAAIRWKNEDAASASYAYS 561

Qy        659 LLIEKAGNSSNATQVVT-DIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCT 717
              :|| | |: || |  | | |    : ||||| ||||:|  |||||  || ||   |||
Db        562 VLILKTGDGSNVTSNFTKDPSI---LIPELIPGVSYTVKILTQVGDGTTSLVPGWNLFCT 618

Qy        718 DPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYR 777
              :|  : || |||||||||||||| || |   |||| | | :|:| | ||:|:|:: || |
Db        619 EPEPVTSFHCEVVPKEPALVLKWACPFGMYTGFELGVRSDSWDNMTRLENCTSDDDTECR 678

Qy        778 TEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKF 837
              ||| |||||||||||| |:|||||| | :| |||||||| |||||||||||||||||||
Db        679 TEVAYLNFSTSYNISIATLSCGKMALPAQNICTTGITDPPTPDGSPNITSVSHNSVKVKF 738

Qy        838 SGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQS 897
              |||||||||||||||||||  ||||||||||:|||:||||||||||| ||||:||
Db        739 SGFEASHGPIKAYAVILTTGEAAQPSADVLKYTYEDFKRGASDTYVTYLIRIEEKGQSQG 798

Qy        898 LSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSF 957
              ||||| |||||||:||||||||:|||||||||||||||:: || |||:| ||||||
Db        799 LSEVLNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYVSF 858

Qy        958 SRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKK 1017
              | ||:|| |||||||||||||||||||| | ||||||||| ||||||||||||||||||
Db        859 SPYSEAVFLPQDPGVICGAVFGCIFGALAITAVGGFIFWRKKRTDAKNNEVSFSQIKPKK 918

Qy       1018 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPY 1077
              |||||||||||||||||||||||||||||||:||| ||| ||:|||||||||||||||
Db        919 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYTAEIAENRGKNRYNNVLPY 978

Qy       1078 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIM 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        979 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIVM 1038

Qy       1138 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQ 1197
              |||||||||||||||||||||||||||||||||:|||||||| |||:| ||||||||
Db       1039 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEVVLPEWTIRDFVVKNMQNSESHPLRQ 1098

Qy       1198 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIY 1257
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1099 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRLIY 1158

Qy       1258 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAM 1317
              |||||||||||||||||||||||||||||||||||||||:|:|||||||||||||||
Db       1159 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTTAM 1218

Qy       1318 TIYENLAPVTTFGKTNGYIA 1337
              |||||| ||: |||||||||
Db       1219 TIYENLEPVSMFGKTNGYIA 1238


RESULT 4
Q62884
ID   Q62884      PRELIMINARY;      PRT;  1216 AA.
AC   Q62884;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Density-enhanced phosphatase-1 precursor (EC 3.1.3.48) (DEP-1)
DE   (Vascular protein tyrosine phosphatase 1).
GN   Name=Dep1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR KYOTO; TISSUE=AORTA;
RX   MEDLINE=96375109; PubMed=8781490;
RA   Borges L.G., Seifert R.A., Grant F.J., Hart C.E., Disteche C.M.,
RA   Edelhoff S., Solca F.F., Lieberman M.A., Lindner V., Fischer E.H.,
RA   Lok S., Bowen-Pope D.F.;
RT   "Cloning and characterization of rat density-enhanced phosphatase-1, a
RT   protein tyrosine phosphatase expressed by vascular cells.";
RL   Circ. Res. 79:570-580(1996).
CC   -!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND/OR SIGNALING EVENTS
CC       INVOLVING CELL-CELL AND CELL-MATRIX CONTACT.
CC   -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN
CC       TYROSINE + PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: HIGH EXPRESSION IN THE CEREBELLUM, BRAIN
CC       CORTEX, AND KIDNEY CORTEX, AND SOMEWHAT LESS ABUNDANT IN SPLEEN
CC       AND LUNG.
CC   -!- SIMILARITY: CONTAINS 8 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 PROTEIN-TYROSINE PHOSPHATASE DOMAIN.
DR   EMBL; U40790; AAB53195.1; -.
DR   HSSP; P18052; 1YFO.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 7.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 8.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 4.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Glycoprotein; Hydrolase; Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29   1216       DENSITY-ENHANCED PHOSPHATASE-1.
FT   DOMAIN       21     28       POLY-LEU.
FT   DOMAIN       29    854       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      695    698       POLY-PRO.
FT   TRANSMEM    855    875       POTENTIAL.
FT   DOMAIN      876   1216       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       39    102       FIBRONECTIN TYPE-III.
FT   DOMAIN      122    225       FIBRONECTIN TYPE-III.
FT   DOMAIN      245    314       FIBRONECTIN TYPE-III.
FT   DOMAIN      333    399       FIBRONECTIN TYPE-III.
FT   DOMAIN      419    486       FIBRONECTIN TYPE-III.
FT   DOMAIN      505    575       FIBRONECTIN TYPE-III.
FT   DOMAIN      597    672       FIBRONECTIN TYPE-III.
FT   DOMAIN      701    787       FIBRONECTIN TYPE-III.
FT   DOMAIN      940   1216       PROTEIN-TYROSINE PHOSPHATASE.
FT   ACT_SITE   1118   1118       BY SIMILARITY.
FT   CARBOHYD     62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     83     83       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     88     88       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    108    108       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    147    147       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    162    162       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    184    184       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    256    256       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    261    261       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    271    271       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    311    311       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    344    344       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    357    357       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    376    376       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    405    405       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    459    459       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    465    465       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    483    483       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    496    496       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    508    508       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    517    517       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    549    549       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    640    640       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    646    646       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    651    651       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    663    663       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    669    669       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    703    703       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    789    789       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    816    816       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1216 AA;  134275 MW;  4D6193B108E53591 CRC64;

  Query Match           64.0%;  Score 4499.5;  DB 2;  Length 1216;
  Best Local Similarity 67.1%;  Pred. No. 1.5e-246;
  Matches  900;  Conservative  106;  Mismatches  205;  Indels  131;  Gaps   10;

Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              |||||||  |||||||||||  |||||| ||::| |
Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLRQGQVVCTG---------------------- 37

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
                                                                 ||
Db         38 ------------------------------------------------------AAPS 41

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||::||: |||:|:|||| ||:| ||||       |   || :|  ||||| |||
Db         42 PVFDVEAVT-SPTSVVLTWKHNDSATSEYKI----NEGNTLRYTVKNQTSFNITGLSPAT 96

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              || |||| |  ||| | |    : ||| |||||:||  || :| |:||| ||||| |: :
Db         97 SYKFSITLGTVNETSGKPTYKNITTEPWPVSDLQVAYIGVTQALLAWSNANGTASYRMQI 156

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
                  ||| :|     |||||||   ::   :  ||:|: |   ||
Db        157 ------VELTTNS--SGGISDLKPGTHKSL--AVQGSNETQHDLWVTE------------ 194

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
                        |  || | :    |||||  |:| |:|  |:|||||:||||||
Db        195 --------------GVSDPPSARDPSLTEVLLTELKPDTQYKVTIYSQAADGTEGQPGNK 240

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              |:|| ||| |: ||||| :::|| || ::||| :::|||:||| :|| |  |:| :||
Db        241 VFKTNPIQVSDIRAVNISDSNMTLTWKSNNNESHASFTYKIYVAGGSDSINETVNETQAV 300

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              | || ||| ||||| | ||   | ||||||:| | |||||||| ||  |||||| |:|:|
Db        301 IRGLSSSTLYNITVLPFLGQTAGIPGFLQVYTSPRPVSDFRVTNVSLREIGLAWRSNDSE 360

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              ||:: |||||:       | : | |:  | ||| | ||| |:|||||| |:|| ||
Db        361 SFEIFITQEGSEKRWNASTGDLSYIVVNLKPGTSYQFEIFPRGPNGTEGPSQTVVGRTDC 420

Qy        541 SAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAITLQGLIPGT 600
              ||| || || |:|||: |:|:: | || | ||||:||::||  |::  | ||: || |||
Db        421 SAVTDIRVVSVSTTEIQLEWQNTDSASGYTYHLVLESENGSIKTNSSQKWITVGGLTPGT 480

Qy        601 LYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYCLL 660
              |||:|| |||| : |: :|  |||||||| |:|:|||||    : |:| | || :||| :|
Db        481 LYNVTIFPEVDQMEGNSSSITQYTRPSNVSYIEVNTNTTVGAIQWKNLDAASASYSYSVL 540

Qy        661 IEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIF-----AQVGDGIKSLEPGRKSF 715
              | |||: || |  | |  |:  |||| |  |:||      :||: :   ||:| |
Db        541 ILKAGDGSNVTSRVRD--IPSVTIPGLIPGVSSEVKIFTKIRNTEVGNEV----PGQKLF 594

Qy        716 CTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTE 775
```

```
Db        595 CMEPAQVDSLHCEVVPKEPALVLKWACPPGMNSGFELGVRSDAWDNMTHLENCTLDNDTE 654

Qy        776 YRTEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKV 835

Db        655 CRTEVTYLNFSTSYNISIATLSCGKMALPTQSTCTTGITDPPPPDGSPNITSVSHNSVKV 714

Qy        836 KFSGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRS 895

Db        715 KFSGFEASHGPIKAYAVLLTTGEAGQPSTDVLKYTYEDFKKGASDTYVTYLIRIEEKGQS 774

Qy        896 QSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYV 955

Db        775 QGLSEALNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYV 834

Qy        956 SFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKP 1015

Db        835 SFSPYSEAVSLPQDPGVICGAVFGCIFGALAIVAVGGFIFWRKKRKDAKNNEVSFSQIKP 894

Qy       1016 KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVL 1075

Db        895 KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYAAEIAENRGKNRYNNVL 954

Qy       1076 PYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1135

Db        955 PYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1014

Qy       1136 IMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPL 1195

Db       1015 VMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEVVLPEWTIRDFVVKNMQSSESHPL 1074

Qy       1196 RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRL 1255

Db       1075 RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRL 1134

Qy       1256 IYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTT 1315

Db       1135 IYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTT 1194

Qy       1316 AMTIYENLAPVTTFGKTNGYIA 1337

Db       1195 AMTIYENLERVSMFGKANGYIA 1216
```

RESULT 5
Q8CIW9
ID Q8CIW9 PRELIMINARY; PRT; 1238 AA.
AC Q8CIW9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protein tyrosine phosphatase receptor-like protein J.
GN Name=Ptprj;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvSl;
RX MEDLINE=22084388; PubMed=12089527; DOI=10.1038/ng903;
RA van Wezel T., Ruivenkamp C.A.L., Zanon C., Stassen A.P.M., Vlcek C.,
RA Csikos T., Tripodis N., Groot P.C., Dauwerse H., van Ommen G.J.B.,
RA Demant P.;
RT "Ptprj is a candidate for the mouse colon-cancer susceptibility locus
RT Scc1 and is frequently deleted in human cancers.";
RL Nat. Genet. 31:295-300(2002).
DR EMBL; AY038877; AAN11409.1; -.
DR EMBL; AY038861; AAN11409.1; JOINED.
DR HSSP; P18052; 1YFO.
DR MGD; MGI:104574; Ptprj.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0007507; P:heart development; IMP.
DR GO; GO:0001570; P:vasculogenesis; IMP.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00060; FN3; 7.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS50853; FN3; 5.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase; Receptor.
SQ SEQUENCE 1238 AA; 136712 MW; DC294E2543995B6C CRC64;

Query Match 63.8%; Score 4483.5; DB 2; Length 1238;
Best Local Similarity 66.9%; Pred. No. 1.2e-245;
Matches 897; Conservative 109; Mismatches 229; Indels 105; Gaps 12;

```
Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Db          1 MKPAARETRTPPRSPGLRWALLPLLLLLRQGQVLCAGAAPNPI---------------- 43

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Db         44 ------------------------------------------------------------ 43

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Db         44 --FDIEAV-VSPTSVLLTWKHNDSGASECR--IENKMESNLTFPVKNQTSCNITGLSPGT 98

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Db         99 SYTFSIISVTTNETLNK----TITTEPWPVSDLHVTSVGVTQARLTWSNANGTASYRMLI 154

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLL-QSNKTKGDPLGTEGGLDASNTERS 299

Db        155 EELTTHS--------SVNISGLKPGTN---NSFAFPESNETQADFAVAEEVPDANGTKRI 203

Qy        300 RAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQA 359

Db        204 PV-TNLSQLHKNSLVS-VDPPSGQDPSLTEILLTDLKPDTQYNATIYSQAANGTEGQPRN 261

Qy        360 IEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA 419

Db        262 KVFKTNSTQVSDVRAMNISASSMTLTWKSNYDGSRTSIVYKIHVAGGTHSVNQTVNKTEA 321

Qy        420 VIPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDA 479

Db        322 IILGLSSSTLYNITVHPFLGQTEGTPGFLQVYTSPDQVSDFRVTNVSTRAIGLAWRSNDS 381

Qy        480 ESFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTV 539

Db        382 KSFEIFIKQDGGEKHRNASTGNQSYMVEDLKPGTSYHFEIIPRGPDGTEGLSSTVNGSTD 441

Qy        540 PSAVFDIHVVYVTTTEMWLDWKSPDGASEYVYHLVIESKHGS-NHTSTYDKAITLQGLIP 598

Db        442 PSAVTDIRVVNISTTEMQLEWQNTDDASGYTYHLVLESKSGSIIRTNSSQKWITVGSLTP 501

Qy        599 GTLYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNFDDASPTYSYC 658

Db        502 GTLYNVTIFPEVDQIQGISNSITQYTRPSSVSHIEVNTTTTTAAIRWKNEDAASASYAYS 561

Qy        659 LLIEKAGNSSNATQVVT-DIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFCT 717

Db        562 VLILKTGDGSNVTSNFTKDPSI---LIPELIPGVSYTVKILTQVGDGTTSLVPGWNLFCT 618

Qy        718 DPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYR 777

Db        619 EPEPVTSFHCEVVPKEPALVLKWACPFGMYTGFELGVRSDSWDNMTRLENCTSDDDTECR 678
```

```
Qy        778 TEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKF 837
              ||| |||||||||||| |:|||||| | :| ||||||||| ||||||||||||||||||||
Db        679 TEVAYLNFSTSYNISIATLSCGKMALPAQNICTTGITDPPTPDGSPNITSVSHNSVKVKF 738

Qy        838 SGFEASHGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQS 897
              |||||||||||||||||||||  ||||||||||:|||:|||||||||||| ||||:||
Db        739 SGFEASHGPIKAYAVILTTGEAAQPSADVLKYTYEDFKRGASDTYVTYLIRIEEKGQSQG 798

Qy        898 LSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSF 957
              ||||| ||||||:|||||||:|||||||||||||||:: || |||:| ||||||
Db        799 LSEVLNYEIDVGNQSTTLGYYNGRLEPLGSYRACVAGFTNITYNLQNDGLINGDESYVSF 858

Qy        958 SRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKK 1017
              | ||:|| |||||||||||||||||||| |  |||||||||| ||||||||||||||||
Db        859 SPYSEAVFLPQDPGVICGAVFGCIFGALAITAVGGFIFWRKKRTDAKNNEVSFSQIKPKK 918

Qy       1018 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPY 1077
              |||||||||||||||||||||||||||||:||| ||| ||:|||||||||||||||||
Db        919 SKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLIGISLPKYTAEIAENRGKNRYNNVLPY 978

Qy       1078 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIM 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        979 DISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIVM 1038

Qy       1138 LTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQ 1197
              |||||||||||||||||||||||||| |||||:|||||||||| |||:  ||||||||
Db       1039 LTKCVEQGRTKCEEYWPSKQAQDYGDITXAMTSEVVLPEWTIRDFVVKNMXNSESHPLRQ 1098

Qy       1198 FHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIY 1257
              |||||||||||||||||||||||||| ||| ||:|||||||||||||||||||||||||
Db       1099 FHFTSWPDHGVPDTTDLLINFRYLVRDDMKQIPPKSPILVHCSAGVGRTGTFIAIDRLIY 1158

Qy       1258 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAM 1317
              |||||||||||||||||||||||||||||||||||||||:|:|||||||||||||||
Db       1159 QIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTTAM 1218

Qy       1318 TIYENLAPVTTFGKTNGYIA 1337
              |||||| ||: |||||||||
Db       1219 TIYENLEPVSMFGKTNGYIA 1238
```

```
RESULT 6
Q6P4H4
ID   Q6P4H4      PRELIMINARY;      PRT;   539 AA.
AC   Q6P4H4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   PTPRJ protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RA   Strausberg R.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC063417; AAH63417.1; -.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   Pfam; PF00041; fn3; 4.
DR   SMART; SM00060; FN3; 4.
DR   PROSITE; PS50853; FN3; 3.
SQ   SEQUENCE   539 AA;  57189 MW;  E970DA55A6B2E38D CRC64;
```

```
  Query Match           39.6%;  Score 2784;  DB 2;  Length 539;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-149;
  Matches  537;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MKPAARERALPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAARERALPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISGLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IPGLRSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAE 480

Qy        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRT 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFQMHITQEGAGNSRVEITTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRT 538
```

```
RESULT 7
Q9W6V5
ID   Q9W6V5      PRELIMINARY;      PRT;  1406 AA.
AC   Q9W6V5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Supporting-cell antigen precursor.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Intestine;
RX   MEDLINE=99296852; PubMed=10366616;
RA   Kruger R.P., Goodyear R.J., Legan P.K., Warchol M., Raphael Y.,
RA   Cotanche D.A., Richardson G.P.;
RT   "The supporting-cell antigen: a receptor-like protein tyrosine
RT   phosphatase expressed in the sensory epithelia of the inner ear.";
RL   J. Neurosci. 19:4815-4827(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Intestine;
RA   Legan P.K.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ238216; CAB41885.2; -.
DR   HSSP; P18052; 1YFO.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 8.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 9.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 9.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase; Signal.
FT   SIGNAL       1     24       Potential.
SQ   SEQUENCE   1406 AA;  154213 MW;  2D609885CC0F367B CRC64;

  Query Match             39.2%;  Score 2752.5;  DB 2;  Length 1406;
  Best Local Similarity   45.7%;  Pred. No. 2.9e-147;
  Matches  624;  Conservative  169;  Mismatches  459;  Indels  112;  Gaps   35;

Qy         47 SVATVATGENGITQISST-------AESF----HKQNGTGTPQVETNTSE--DGESSGAN 93
              :|  : |   |:| :: |       ::|:      :||:    :   ||:  | | :|
Db         82 AVLDLKTEYVGVTSVNLTWTVNDTASDSYTYRIEVRNGS---SINNKTSDITDAEITGLI 138

Qy         94 DS-------LRTPEQGSNGTDGASQKTPSSTGPSPVFDIKAVSISPTNVILTWKSNDTAA 146
                       |    :|||       | |||| |:||  :  |:| |||  ||||:
Db        139 PGTLYTFTVFAVAADGQTAGEGAS--ISLYTKPSPVLDLKAEYVGVTSVNLTWTVNDTAS 196

Qy        147 SEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSYVFSI-TPGIGNETWGDPRVIKVIT 205
              : | | ::   ::              |||| | | | | :      | |  |   | : |
Db        197 ASYTYRIEVTSDSSIDSLTSSVTMAEITGLIPGTLYSFKVFAVAADNRTEADGASISLYT 256

Qy        206 EPIPVSDLRVALTGVRKAALSWSNGNGTA----SCRVLLESIGSHEELTQDSRLQVNISD 261
              :| || ||:    ||    |:|:  ||||    : |: : :  |    | :   :: |:
Db        257 KPSPVLDLKAEYVGVTSVNLTWT-VNGTALTAYTYRIEVRNATSIRNETSNIN-KIEITG 314

Qy        262 LKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSRAGSPTAPVHDE-------SLV 314
              | ||  ||   :   | |       | ||  :  :    ||   | |       :|
Db        315 LIPGTSYNFKVFATPVNNTT-----EEEGLSLNLYTKP---SPVLRVVTEYVGVTSVNLT 366

Qy        315 GPVDPS----------------SGQQSRDTEVLLVGLEPGTRYNATVYSQAANG-TEGQ 356
                || :               : : |  |:  : || ||| |  ||:: ||:| | |:
Db        367 WMVDDTASDSYTYRIEVRNGSSINNKTSDITDAEITGLIPGTLYTFTVFAVAADGQTAGE 426

Qy        357 PQAIEFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSE 416
              :|    |    | |:  |   :  ||: | | |:|  |:| |||:| |  ::   :|  :
Db        427 GASISLYTKPSPVLDLKAEYVGVTSVNLTWTVNDTASAS-YTYRIEVTSDSSIDSLTSNV 485

Qy        417 PRAVIPGLRSSTFYNITVCPVLGD--IEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAW 474
              | | ||   | || ||  |  |    |    | ::| | ||:| :      | : | |
Db        486 TMAEITGLIPGTLYNFTVFAVAADNRTEADGAFTSLYTKPTPVTDLKAE-HGVTSVSLNW 544

Qy        475 SSHDAESFQMHITQEGAGNSRVEITTNQSI----------IIGGLFPGTKYCFEIVPKGP 524
                :|        |   :   |:|:    |:           | || ||| | | :
```

```
Db        545 MVND--------TASDSYTYRIEVRNGHSVNNKTSNIPETEITGLNPGTLYTFTVFAVAA 596

Qy        525 NG-TEGASRTVCNRTVPSAVFDIHVVYVTTTEMWLDWKSPD-GASEYVYHLVIESKHGS- 581
              :| |||   ::   | | ||  :    ||  | : | |    |  :: | |   || ::||
Db        597 DGETEGEGASISVYTKPRAVLHLKTEYVGVTSVNLTWTVNDTDSASYTYR--IEVRNGSS 654

Qy        582 --NHTSTYDKAITLQGLIPGTLYNITISPEV--DHVWGDPNSTAQYTRPSNVSNIDVSTN 637
                | ||    |  : || |||||  |:           |:  | : ||:|| | |:
Db        655 INNKTSDITDA-EITGLDPGTLYIFTVFAVAADGQTAGEGASISLYTKPSMVLNLKAEYV 713

Qy        638 T-TAATLSWQNFDDASPTYSYCLLIEKAGNS-SNATQVVTDIGITDATVTELIPGSSYTV 695
              | |:  |:|   |  | :|:|   || |  |  | |       :| : || ||||:||
Db        714 TMTSVNLTWMVNDAESASYTY--RIEVAHESLINETM----SNVTKSIVTYLIPGTSYNF 767

Qy        696 EIFAQVGDGIKSLEPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVS 755
              :||  |    |   |  | |:|: :| || |    | ||| || |  :||::|:
Db        768 TVFAIAADNQTEGEGASISQNTVPSSVNAFQCEAVANMSYLTLKWNCPYGGYSGFDIEIF 827

Qy        756 SGAWNNATHLESCSSENGTE-YRTEVTYLNFSTSYNISITTVSCGKMAAPTRNTCTTGIT 814
              :| |      : |  |   | ::||   |::  :| :|:|||| |  : | :  | | ||
Db        828 NGTWTKKQQSQFCGREGSEEIFKTEP--LDYYKTYTVSVTTVSDGLTSLPVQKICKTSIT 885

Qy        815 DPPPPDGSPNITSVSHNSVKVKFSGFEASHGPIKAYAV-ILTTGEAGHPSADVLKYTYDD 873
              ||| |: :| : :|||||: |:|  ||: :||:|||| |:|  |   ||   | |||:|
Db        886 DPPVPNKAPLVKAVSHNSLSVEFPDFESVNGPLKAYAVMIVTEAEGCLPSKSDLDYTYND 945

Qy        874 FKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLGYYNGKLEPLGSYRACVA 933
              ||:  : |||||:|  ||   | | |:     :|||  :|  || || | || |||| ||
Db        946 FKQKMTATYVTYVIDVEEIS-SSSHSQNGHNIVDVGKGNTMYGYENGPLIPLHSYRASVA 1004

Qy        934 GFTNITFHPQNKGLIDGAESYVSFSRYSDAVSLPQDPGVICGAVFGCIFGALVIVTVGGF 993
              ||||| |   ||  | | :||||||  |: ||||||||| ||| ||:   | :| :||:
Db       1005 GFTNINFTVANK--IMGEQSYVSFSPCSEVVSLPQDPGVIAGAVIGCLLAILAVVAIGGY 1062

Qy        994 IFWRKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGI 1053
              ||||::||| :| ||||| || ||||:|:|||||:||||||||||||||||||:||  |:
Db       1063 IFWRRRRKDKRNTEVSFSPIKIKKSKMIKVENFESYFKKQQADSNCGFAEEYEELKSAGV 1122

Qy       1054 SQPKYAAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQ 1113
               |||:|||:|||||||||||||||||||||||  : :|||||||||||| ||| ||| |
Db       1123 HQPKFAAEIAENRGKNRYNNVLPYDISRVKLSNPSCTTDDYINANYMPGYSSKKAFIAAQ 1182

Qy       1114 GPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIV 1173
              ||||||::|||||:||||:|:|:|||||||| |||||:||| ||:: |||| | | ||:|
Db       1183 GPLPNTIEDFWRMIWEKNIYSIVMLTKCVEQARTKCEQYWPDKQSKSYGDIIVTMVSEVV 1242

Qy       1174 LPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPES 1233
              |||||||| |:|  | ||| :|||||||||||||:|||||||:|| :|  |:| :|
Db       1243 LPEWTIRDFNVENADTMESHTVRQFHFTSWPDHGVPETTDLLINFRHLVHEYSSQNPIDS 1302

Qy       1234 PILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQ 1293
              |||||||||||||||||||||| ||| ||||||||:||||||||||||||||||||||
Db       1303 PILVHCSAGVGRTGTFIAIDRLIQQIEMENTVDVYGVVYDLRMHRPLMVQTEDQYVFLNQ 1362

Qy       1294 CVLDIVRSQKDSKVDLIYQNTTAMTIYENLAPVTTFGKTNGYIA 1337
              ||:||:|||: | |||||||||| ||||  |  ||| ||| |
Db       1363 CVMDIIRSQKEKKTDLIYQNTTAMAIYENFTPGPAFGKANGYHA 1406
```

```
RESULT 8
Q9NPR5
ID   Q9NPR5      PRELIMINARY;      PRT;   425 AA.
AC   Q9NPR5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   PTPRJ, protein tyrosine phosphatase receptor J, eta (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA   Lehrach H., Poustka A., Lundeberg J.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Andreu N., Estivill X., Escarceller M., Sumoy L.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL359057; CAB94390.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   Pfam; PF00041; fn3; 2.
DR   SMART; SM00060; FN3; 2.
DR   PROSITE; PS50853; FN3; 1.
KW   Receptor.
FT   NON_TER     425    425
SQ   SEQUENCE   425 AA;  45078 MW;  BD1ACA5D345B7027 CRC64;

  Query Match             31.1%;  Score 2184;  DB 2;  Length 425;
  Best Local Similarity   99.8%;  Pred. No. 8.7e-116;
  Matches  424;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKPAAREARLPPRSPGLRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVATGENGITQ 60

Qy         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ISSTAESFHKQNGTGTPQVETNTSEDGESSGANDSLRTPEQGSNGTDGASQKTPSSTGPS 120

Qy        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPAT 180

Qy        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRVLL 240

Qy        241 ESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 ESIGSHEELTQDSRLQVNISGLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSR 300

Qy        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAI 360

Qy        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EFRTNAIQVFDVTAVNISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRAV 420

Qy        421 IPGLR 425
              |||||
Db        421 IPGLR 425


RESULT 9
Q61373
ID   Q61373      PRELIMINARY;      PRT;   361 AA.
AC   Q61373;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Protein tyrosine phosphatase (Fragment).
GN   Name=Ptprj; Synonyms=PTP-RL9;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57/BL6; TISSUE=Liver;
RA   Higashitsuji H., Arii S., Furutani M., Imamura M., Kaneko Y.,
RA   Takenawa J., Nakayama H., Fujita J.;
RT   "Enhanced expression of multiple protein tyrosine phosphatases in the
RT   regenerating mouse liver : Isolation of PTP-RL 10, a novel
RT   cytoplasmic-type phosphatase with sequence homology to cytoskeletal
RT   protein 4.1.";
RL   Oncogene 0:0-0(1995).
DR   EMBL; D49393; BAA08386.1; -.
DR   HSSP; P18052; 1YFO.
DR   MGD; MGI:104574; Ptprj.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   GO; GO:0007507; P:heart development; IMP.
DR   GO; GO:0001570; P:vasculogenesis; IMP.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase.
FT   NON_TER       1      1
SQ   SEQUENCE   361 AA;  41726 MW;  B51462E8D11023AA CRC64;

  Query Match             26.3%;  Score 1847;  DB 2;  Length 361;
  Best Local Similarity   94.7%;  Pred. No. 9.2e-97;
  Matches  342;  Conservative    8;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        977 VFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQAD 1036
              |||||||| |  |||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 VFGCIFGALAITAVGGFIFWRKKRTDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQAD 60

Qy       1037 SNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYIN 1096
              ||||||||||||||:||| ||| ||:|||||||||||||||||||||||||||||||||||
Db         61 SNCGFAEEYEDLKLIGISLPKYTAEIAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYIN 120

Qy       1097 ANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSK 1156
              |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        121 ANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIVMLTKCVEQGRTKCEEYWPSK 180

Qy       1157 QAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLI 1216
              |||||||||||||:|||||||||| |||:| ||||||||||||||||||||||||||||||
Db        181 QAQDYGDITVAMTSEVVLPEWTIRDFVVKNMQNSESHPLRQFHFTSWPDHGVPDTTDLLI 240

Qy       1217 NFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRM 1276
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NFRYLVRDYMKQIPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDLRM 300

Qy       1277 HRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIYENLAPVTTFGKTNGYI 1336
              |||||||||||||||||||:|:|||||||||||||||||| ||: ||||||||
Db        301 HRPLMVQTEDQYVFLNQCVLDIIRAQKDSKVDLIYQNTTAMTIYENLEPVSMFGKTNGYI 360

Qy       1337 A 1337
              |
Db        361 A 361


RESULT 10
Q8CIW2
ID   Q8CIW2      PRELIMINARY;      PRT;  1998 AA.
AC   Q8CIW2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Vascular endothelial protein tyrosine phosphatase.
GN   Name=Ptprb;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster;
```

```
RX   MEDLINE=22220060; PubMed=12234928; DOI=10.1093/emboj/cdf497;
RA   Nawroth R., Poell G., Ranft A., Kloep S., Samulowitz U., Fachinger G.,
RA   Golding M., Shima D.T., Deutsch U., Vestweber D.;
RT   "VE-PTP and VE-cadherin ectodomains interact to facilitate regulation
RT   of phosphorylation and cell contacts.";
RL   EMBO J. 21:4885-4895(2002).
DR   EMBL; AY077755; AAL75813.1; -.
DR   HSSP; P18052; 1YFO.
DR   MGD; MGI:97809; Ptprb.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 15.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 16.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 15.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase.
SQ   SEQUENCE   1998 AA;  224474 MW;  B37EA6501EF03E5D CRC64;

  Query Match             17.9%;  Score 1259.5;  DB 2;  Length 1998;
  Best Local Similarity   28.8%;  Pred. No. 3e-62;
  Matches  378;  Conservative  216;  Mismatches  500;  Indels  217;  Gaps   45;

Qy        104 NGTDGASQKTPSSTGPSPVFDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTI 163
Db        805 SGQYEASEQGTGRTIPEPVKDLTLLNRSTEDLHVTWSRANGDVDQYE--VQLLFNDMKVF 862

Qy        164 TVVHQPWCNITGLRPATSYVF-SITPG-----IGNETWGDPR---VIKVITEPIPVSDLR 214
Db        863 PHIHL-------VNTATEYKFTALTPGRHYKILVLTISGDVQQSAFIEGLPVPSTVKNIH 915

Qy        215 VALTGVR-KAALSWSNGNGTASCRVLLESIGSHEELTQDSRLQVNISDLKPGVQYNINPY 273
Db        916 ISANGATDRLMVTWSPGGG-----------------------------------DVDSY 939

Qy        274 LLQSNKTKGDPLGTEGGLDASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLV 333
Db        940 VVSAFR-----------------QDEKVDSQTIPKH-----------------ASEHTFH 965

Qy        334 GLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVTAVN-ISATSLTLIWKVSDNE 392
Db        966 RLEAGAKYRIAIVS-VSGSLRNQIDALG-QTVPASVQGVVAANAYSSNSLTVSWQKALGV 1023

Qy        393 SSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGLRSSTFYNITVCPVLGDIEGT 444
Db       1024 AE---RYDILLLNENGLLLSNVSEPATARQHKFEDLTPGKK----YKMQILTVSGGLFSK 1076

Qy        445 PGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQMHI---TQEGAGNSRVEI-TT 500
Db       1077 ESQAEGRTVPAAVTNLRITENSSRYLSFGWTASEGELSWYNIFLYNPDRTLQERAQVDPL 1136

Qy        501 NQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHVVYVTTTE-MWLD 559
Db       1137 VQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVPAAVNHLKGSHRNTTDSLWFS 1194

Qy        560 WKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGLIPG---TLYNITISPEVDHV 613
Db       1195 WSPASGDFDF-YELILYNPNGTKKENWKEKDVTEWRFQGLVPGRKYTLYVVTHSGDLSNK 1253

Qy        614 WGDPNSTAQYTRPSNVSNIDVSTNTTAATL----SWQNFDDAS----PTYSYCLLIEKAG 665
Db       1254 VTGEGRTAP-SPPSLLSFADVANTSLAITWKGPPDWTDYNDFELQWFPGDALTIFNPYSS 1312

Qy        666 NSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSL-EPGRKSFCTDPASMAS 724

Db       1313 RKSEG-----------RIVYGLHPGRSYQFSVKTVSGDSWKTYSKPISGSVRTKPDKIQN 1361

Qy        725 FDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEVTYLN 784
Db       1362 LHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CRKMDTQEIEFSRKLEKEKSLLN 1412

Qy        785 F-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNI------TSVSHNSV 833
Db       1413 IMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPQP--PPHIRVNEKDVLISKSSI 1470

Qy        834 K--VKFSGPEASHGPIKAYAVILTTG--------EAGHPSADVLKYTYDDFKKGASDTYV 883
Db       1471 NFTVNCSWFSDTNGAVKYFAVVVREADSMDELKPEQQHPLPSYLEYRHNASIRVYQTNY- 1529

Qy        884 TYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG---------YYNGKLEPLGSYRACVAG 934
Db       1530 -FASKCAESPDSSSKS----FNIKLGAEMDSLGGKCDPSQQKFCDGPLKPHTAYRISIRA 1584

Qy        935 FTNITFHPQNKGLIDGAESYVSFSRYSDAVSLPQDP--GVICGAVFGCIFGALVIVTVGG 992
Db       1585 FTQLF----DEDLKEFTKPLYSDTFFSMPITTESEPLFGVIEGVSAGLFLIGMLVALVAF 1640

Qy        993 FIFWRK--KRKDAKNNEVSFSQIKP-------------KKSKLIRVENPEAYFKKQQADS 1037
Db       1641 FICRQKASHSRERPSARLSIRRDRPLSVHLNLGQKGNRKTSCPIKINQFEGHFMKLQADS 1700

Qy       1038 NCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVKLS-VQTHSTDDYIN 1096
Db       1701 NYLLSKEYEDLKDVGRSQSCDIALLPENRGKNRYNNILPYDASRVKLSNVDDDPCSDYIN 1760

Qy       1097 ANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSK 1156
Db       1761 ASYIPGNNFRREYIATQGPLPGTKDDFWKMAWEQNVHNIVMVTQCVEKGRVKCDHYWPAD 1820

Qy       1157 Q-AQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPL-RQFHFTSWPDHGVPDTTDL 1214
Db       1821 QDPLYYGDLILQMVSESVLPEWTIREFKICSEEQLDAHRLIRHFHYTVWPDHGVPETTQS 1880

Qy       1215 LINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIENENTVDVYGIVYDL 1274
Db       1881 LIQFVRTVRDYINRSPGAGPTVVHCSAGVGRTGTFVALDRILQQLDSKDSVDIYGAVHDL 1940

Qy       1275 RMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIYENLAP 1325
Db       1941 RLHRVHMVQTECQYVYLHQCVRDVLRAKKLRN----EQENPLFPIYENVNP 1987


RESULT 11
PTPB_HUMAN
ID   PTPB_HUMAN     STANDARD;      PRT;  1997 AA.
AC   P23467;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Protein-tyrosine phosphatase beta precursor (EC 3.1.3.48) (R-PTP-
DE   beta).
GN   Name=PTPRB; Synonyms=PTPB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=91006018; PubMed=2170109;
RA   Krueger N.X., Streuli M., Saito H.;
RT   "Structural diversity and evolution of human receptor-like protein
RT   tyrosine phosphatases.";
RL   EMBO J. 9:3241-3252(1990).
CC   -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC       tyrosine + phosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
```

```
CC   -!- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC       Receptor class 3 subfamily.
CC   -!- SIMILARITY: Contains 17 fibronectin type III domains.
CC   -!- SIMILARITY: Contains 1 protein-tyrosine phosphatase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54131; CAA38066.1; -.
DR   PIR; S12050; S12050.
DR   HSSP; P18052; 1YFO.
DR   Genew; HGNC:9665; PTPRB.
DR   MIM; 176882; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005001; F:transmembrane receptor protein tyrosine pho. . .; TAS.
DR   GO; GO:0006796; P:phosphate metabolism; TAS.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 16.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 17.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 17.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Glycoprotein; Hydrolase; Phosphorylation; Protein phosphatase; Repeat;
KW   Signal; Transmembrane.
FT   SIGNAL        1     22       Potential.
FT   CHAIN        23   1997       Protein-tyrosine phosphatase beta.
FT   DOMAIN       23   1621       Extracellular (Potential).
FT   TRANSMEM   1622   1642       Potential.
FT   DOMAIN     1643   1997       Cytoplasmic (Potential).
FT   DOMAIN       23    108       Fibronectin type-III 1.
FT   DOMAIN      110    201       Fibronectin type-III 2.
FT   DOMAIN      203    288       Fibronectin type-III 3.
FT   DOMAIN      289    375       Fibronectin type-III 4.
FT   DOMAIN      377    465       Fibronectin type-III 5.
FT   DOMAIN      467    552       Fibronectin type-III 6.
FT   DOMAIN      554    641       Fibronectin type-III 7.
FT   DOMAIN      643    730       Fibronectin type-III 8.
FT   DOMAIN      731    817       Fibronectin type-III 9.
FT   DOMAIN      819    906       Fibronectin type-III 10.
FT   DOMAIN      907    993       Fibronectin type-III 11.
FT   DOMAIN      995   1083       Fibronectin type-III 12.
FT   DOMAIN     1085   1172       Fibronectin type-III 13.
FT   DOMAIN     1173   1260       Fibronectin type-III 14.
FT   DOMAIN     1261   1353       Fibronectin type-III 15.
FT   DOMAIN     1355   1443       Fibronectin type-III 16.
FT   DOMAIN     1448   1550       Fibronectin type-III 17.
FT   DOMAIN     1722   1997       Protein-tyrosine phosphatase.
FT   ACT_SITE   1904   1904       Phosphocysteine intermediate (By
FT                                similarity).
FT   CARBOHYD     28     28       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     53     53       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     75     75       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    172    172       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    198    198       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    267    267       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    321    321       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    414    414       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    421    421       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    479    479       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    544    544       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    574    574       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    598    598       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    652    652       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    721    721       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    829    829       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1040   1040       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1096   1096       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1163   1163       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1185   1185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1212   1212       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1274   1274       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1367   1367       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1470   1470       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1474   1474       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1518   1518       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   1997 AA;  224267 MW;  691E99BA7A1515DD CRC64;

  Query Match           17.7%;  Score 1242;  DB 1;  Length 1997;
  Best Local Similarity 29.6%;  Pred. No. 3e-61;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy       123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
Db       759 FDHYEVTIKNKNNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 796

Qy       183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
Db       797 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 844

Qy       239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
Db       845 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 901

Qy       281 KG-----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
Db       902 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 957

Qy       314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
Db       958 ------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 1002

Qy       374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
Db       1003 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1059

Qy       425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
Db       1060 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1115

Qy       485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
Db       1116 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1173

Qy       541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
Db       1174 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1232

Qy       597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
Db       1233 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1289

Qy       648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
Db       1290 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQFNVKTVSGDS 1339

Qy       705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
Db       1340 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1390

Qy       764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
Db       1391 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1450
```

```
Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
              |   |:|         :| :|:    |  | |  ::| :| : |::           |  ||
Db       1451 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1508

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
                |:| ::   :      |  :  :  |   | | |    : | :| |  :||
Db       1509 LPSYLEYRHNASIRVYQTNY--PASKCAENPNSNSKS----FNIKLGAEMESLGGKRDPT 1562

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
                : :| |:|  :||  :  ||          | |  |    |::  |||   |||
Db       1563 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1611

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
               :|  : ||: |   |  :|   | |   :  |:|    ::  :  :|  : :|
Db       1612 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1669

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
                     | |  |::  || :| | |||||   ::|||:|| || :|    | | ||||
Db       1670 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1729

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
              ||||||:|||| :||||| |      |||||:|:|| : ::::| |||||| |  |||:|
Db       1730 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1789

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
              |||:||: |:|:|:|||:|| ||: |||: | :  |||: : | || |||||||||:| :
Db       1790 VWEQNVHNIVMVTQCVEKGRVKCDHYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1849

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
                :  ::| | | ||:| |||||||:||  || |   ||||: :||   | :||||||
Db       1850 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1909

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
              |||||||:||:: |::::::||:|| |:|||:||  ||||| |||:|:|||  |::|::|
Db       1910 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1969

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325
                      |      ||||: |
Db       1970 RS----EQENPLFPIYENVNP 1986


RESULT 12
Q6MZF6
ID   Q6MZF6      PRELIMINARY;      PRT;  1956 AA.
AC   Q6MZF6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686E13109 (Fragment).
GN   Name=DKFZp686E13109;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Human fetal kidney;
RG   The German Human cDNA Consortium;
RA   Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA   Han M., Wiemann S.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX648771; CAE46198.1; -.
DR   HSSP; P18031; 1KAV.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003595; PTPc_motif.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 14.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 16.
DR   SMART; SM00194; PTPc; 1.
DR   SMART; SM00404; PTPc_motif; 1.
DR   PROSITE; PS50853; FN3; 15.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase; Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   1956 AA;  219892 MW;  08EC5C8FAF17221E CRC64;

  Query Match           17.6%;  Score 1239;  DB 2;  Length 1956;
  Best Local Similarity 29.6%;  Pred. No. 4.2e-61;
  Matches  397;  Conservative  190;  Mismatches  503;  Indels  251;  Gaps   53;

Qy        123 FDIKAVSISPTNVILTWKSNDTAASEYKYVVKHKMENEKTITVVHQPWCNITGLRPATSY 182
              ||   |:|   |  :  ||               | |||          |    | |  |
Db        718 FDHYEVTIKNKNNFIQTKSIP-----------KSENE----------CVFVQLVPGRLY 755

Qy        183 VFSITPGIG----NETWGDPRVIKVITEPIPVSDLRVALTGVRKAALSWSNGNGTASCRV 238
                ::|   |    ||  |: | |     | || || :         ::||  ||
Db        756 SVTVTTKSGQYEANEQ-GNGRTI-----PEPVKDLTLRNRSTEDLHVTWSGANG------ 803

Qy        239 LLESIGSHE--ELTQDSRL----------QVNISDLKPGVQY-----NINPYLLQSNKT 280
                  :  :|   | | ::          :  : | || ||      |:  : ||
Db        804 ---DVDQYEIQLLFNDMKVFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQSAFI 860

Qy        281 KG-----------DPLG---------TEGGLDASN------TERSRAGSPTAPVHDESL 313
              :|           | |         | || | :          :  | | | |
Db        861 EGFTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDSYTVSAFRHSQKVDSQTIPKH---- 916

Qy        314 VGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVT 373
                          |     || | :|   : |   :|:      :  ||   | |
Db        917 -------------VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTVPASVQGVI 961

Qy        374 AVN-ISATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGL 424
              | |  |: || : |: :   :     | | :  |        | |||          : ||
Db        962 ADNAYSSYSLIVSWQKAAGVAE---RYDILLLTENGILLRNTSEPATTKQHKFEDLTPGK 1018

Qy        425 RSSTFYNITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQM 484
              :   | | :  | | :      :  | |  |:| |:|  ||  :   |:: : |
Db       1019 K----YKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWY 1074

Qy        485 HI---TQEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVP 540
              :|    :|    | ::    ||     |  |  |   ||    :|       :  ||||
Db       1075 NIFLYNPDGNLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVP 1132

Qy        541 SAVFDIHVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGL 596
              ::|  :      ||: :| :|     |  :: | |:: : :|:    :  || :|   |||
Db       1133 ASVSHLRGSNRNTTDSLWFNWSPASGDFDF-YELILYNPNGTKKENWKDKDLTEWRFQGL 1191

Qy        597 IPGTLYNITISPEVDHVWGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL----SWQN 647
              :||  | : :   | |     |    :|     || :|  |::  : | |      | :
Db       1192 VPGRKYVLWV---VTHSGDLSNKVTAESRTAPSPPSLMSFADIANTSLAITWKGPPDWTD 1248

Qy        648 FDDASPTY---SYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDG 704
              ::|    :       :     |  :  ::|           | || ||   :    ||
Db       1249 YNDFELQWLPRDALTVFNPYNNRKSEGRIVYG----------LRPGRSYQSNVKTVSGDS 1298

Qy        705 IKSL-EPGRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNAT 763
              |:  :|   |  | |  : :  |        |:   | |      |: :|
Db       1299 WKTYSKPIFGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CR 1349

Qy        764 HLESCSSENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPP 818
               :::   |   :   | : ||          | :||   | |  :    ::  | |  |||
Db       1350 KMDTQEVEFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPP 1409

Qy        819 PDGSPNI------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHP 862
              |   |:|         :| :|:    |  | |  ::| :| : |::           |  ||
Db       1410 P--PPHIRVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFTVVVREADGSDELKPEQQHP 1467

Qy        863 SADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG------ 916
```

```
                 |:| ::   :     |  :  :  |   | | |    : | :| |  :||
Db       1468 LPSYLEYRHNASIRVVYQTNY--FASKCVENPNSNSKS----FNIKLGAEMESLGGKCDPT 1521

Qy        917 ---YYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSR--YSDA-VSLP--- 967
              : :| |:|  :|| :  ||          | |  |    |::  |||   |||
Db       1522 QQKFCDGPLKPHTAYRISIRAFTQ---------LFD--EDLKEFTKPLYSDTFFSLPITT 1570

Qy        968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNNEVSFSQIKP----- 1015
               :|  : ||: |   |  :| - | |   :  |:|    ::  :  :|  : :|
Db       1571 ESEP--LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRRDRPLSVHL 1628

Qy       1016 --------KKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRG 1067
                     | |  |::  || :| | |||||   ::|||:|| || :|    | | ||||
Db       1629 NLGQKGNRKTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDIALLPENRG 1688

Qy       1068 KNRYNNVLPYDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRM 1126
              ||||||:|||| :||||| |      |||||:|:|| : ::::| |||||| |  |||:|
Db       1689 KNRYNNILPYDATRVKLSNVDDDPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDFWKM 1748

Qy       1127 VWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVK 1185
              |||:||: |:|:|:|||:|| ||: |||: | :  |||: : | || ||||||||:| :
Db       1749 VWEQNVHNIVMVTQCVEKGRVKCDRYWPADQDSLYYGDLILQMLSESVLPEWTIREFKIC 1808

Qy       1186 NIQTSESHPL-RQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVG 1244
                :  ::| | |  ||:| |||||||:||  || |   ||||: :||   | :||||||
Db       1809 GEEQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVG 1868

Qy       1245 RTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKD 1304
              |||||||:||:: |:::::::||:|| |:|||:||  ||||| |||:|:||| |::|::|
Db       1869 RTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRARKL 1928

Qy       1305 SKVDLIYQNTTAMTIYENLAP 1325
                    |     ||||: |
Db       1929 RS----EQENPLFPIYENVNP 1945
```

```
RESULT 13
Q80VN7
ID   Q80VN7      PRELIMINARY;      PRT;  1102 AA.
AC   Q80VN7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Ptprb protein (Fragment).
GN   Name=Ptprb;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

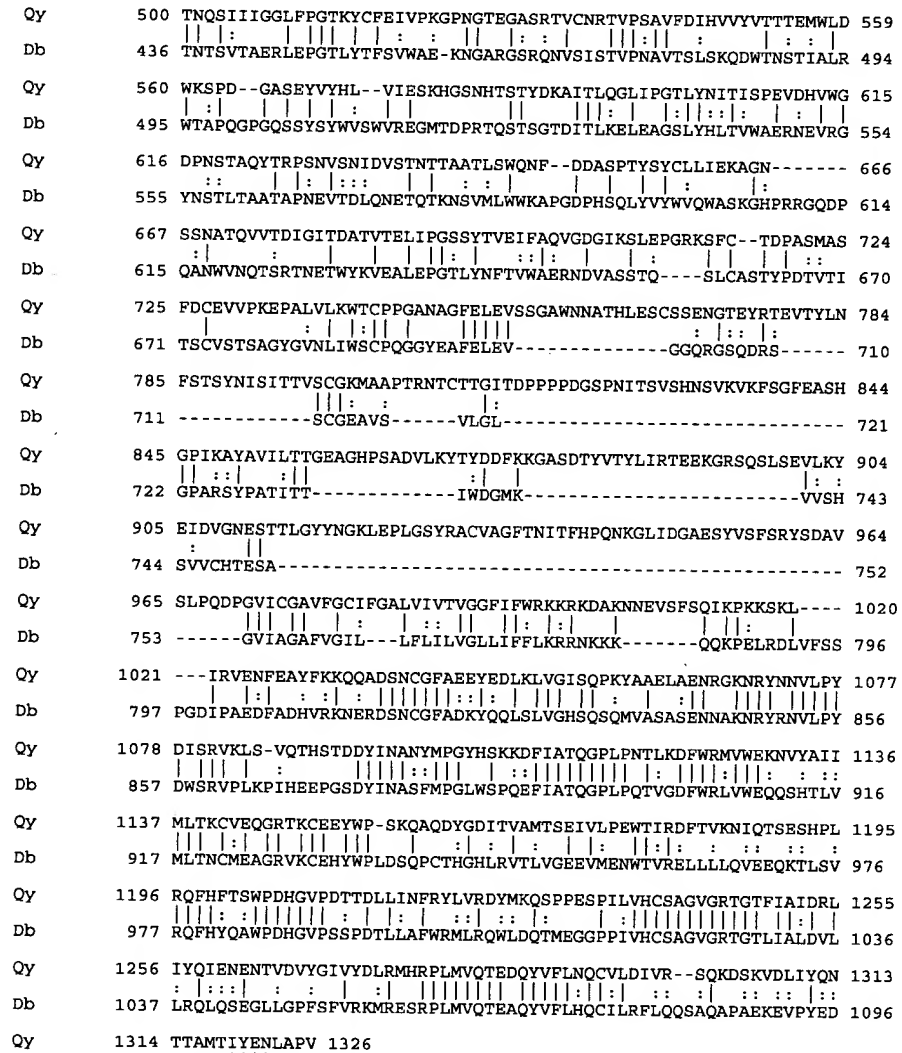
```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC047086; AAH47086.1; -.
DR   HSSP; P18052; 1YFO.
DR   MGD; MGI:97809; Ptprb.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 6.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 6.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 6.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase.
FT   NON_TER       1      1
SQ   SEQUENCE   1102 AA;  124677 MW;  8281DF7F29BF9148 CRC64;

  Query Match           17.3%;  Score 1219.5;  DB 2;  Length 1102;
  Best Local Similarity 29.4%;  Pred. No. 2.2e-60;
  Matches  354;  Conservative  197;  Mismatches  455;  Indels  199;  Gaps   40;

Qy        201 IKVITEPIPVSDLRVALTGVR-KAALSWSNGNGTASCRVLLESIGSHEELTQDSRLQVNI 259
              |: :| |  | :: ::  |   :  ::|| | |
Db          6 IEGLTVPSTVKNIHISANGATDRLMVTWSPGGG-------------------------- 38

Qy        260 SDLKPGVQYNINPYLLQSNKTKGDPLGTEGGLDASNTERSRAGSPTAPVHDESLVGPVDP 319
                    ::: |:: : :                  :  :  | | | |
Db         39 ---------DVDSYVVSAFR-----------------QDEKVDSQTIPKH---------- 62

Qy        320 SSGQQSRDTEVLLVGLEPGTRYNATVYSQAANGTEGQPQAIEFRTNAIQVFDVTAVN-IS 378
                  :|      || | :|   : |  :      |  |:  :|     |  | | |  |
Db         63 -------ASEHTFHRLEAGAKYRIAIVS-VSGSLRNQIDALG-QTVPASVQGVVAANAYS 113

Qy        379 ATSLTLIWKVSDNESSSNYTYKIHVAGETDSSNLNVSEPRA--------VIPGLRSSTFY 430
              : |||: |: :   :      | | :  |      |||||           : || :   |
Db        114 SNSLTVSWQKALGVAE---RYDILLLNENGLLLSNVSEPATARQHKFEDLTPGKK----Y 166

Qy        431 NITVCPVLGDIEGTPGFLQVHTPPVPVSDFRVTVVSTTEIGLAWSSHDAESFQMHI---T 487
               :  |  |  :        :  | |  |:: |:|  |:  :   |:: : |    :|
Db        167 KMQILTVSGGLFSKESQAEGRTVPAAVTNLRITENSSRYLSFGWTASEGELSWYNIFLYN 226

Qy        488 QEGAGNSRVEI-TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDI 546
               :    | ::    ||     |  |  |   ||    :|       :  ||||:||  :
Db        227 PDRTLQERAQVDPLVQSFSFQNLLQGRMYKMVIVTH--SGELSNESFIFGRTVPAAVNHL 284

Qy        547 HVVYVTTTE-MWLDWKSPDGASEYVYHLVIESKHGSNHTSTYDKAIT---LQGLIPG--- 599
               :  ||: :|  |     |  :: | |:: : :|:   :  :| :|    |||:||
Db        285 KGSHRNTTDSLWFSWSPASGDFDF-YELILYNPNGTKKENWKEKDVTEWRFQGLVPGRKY 343

Qy        600 TLYNITISPEVDHVWGDPNSTAQYTRPSNVSNIDVSTNTTAATL----SWQNFDDAS--- 652
              ||| :| | :: :      ||  : || :|  ||:  : | |      | :::|
Db        344 TLYVVTHSGDLSNKVTGEGRTAP-SPPSLLSFADVANTSLAITWKGPPDWTDYNDFELQW 402

Qy        653 -PTYSYCLLIEKAGNSSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSL-EP 710
              |  :  :   :  |                |  | || ||   :   ||  |:  :|
Db        403 FPGDALTIFNPYSSRKSEG-----------RIVYGLHPGRSYQFSVKTVSGDSWKTYSKP 451

Qy        711 GRKSFCTDPASMASFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSS 770
              |  | |  : :  |        |:   |  |      |: :|           :::
Db        452 ISGSVRTKPDKIQNLHCR-PQNSTAIACSWIPPDSDFDGYSIE--------CRKMDTQEI 502

Qy        771 ENGTEYRTEVTYLNF-----STSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNI 825
```

```
Db        503 EFSRKLEKEKSLLNIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTITMIDRPPQP--PPHI 560

Qy        826 ------TSVSHNSVK--VKFSGFEASHGPIKAYAVILTTG--------EAGHPSADVLKY 869

Db        561 RVNEKDVLISKSSINFTVNCSWFSDTNGAVKYFAVVVREADSMDELKPEQQHPLPSYLEY 620

Qy        870 TYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLKYEIDVGNESTTLG---------YYNG 920

Db        621 RHNASIRVYQTNY--FASKCAESPDSSSKS----FNIKLGAEMDSLGGKCDPSQQKFCDG 674

Qy        921 KLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRYSDAVSLPQDP--GVICGAVF 978

Db        675 PLKPHTAYRISIRAFTQLF----DEDLKEFTKPLYSDTFFSMPITTESEPLFGVIEGVSA 730

Qy        979 GCIFGALVIVTVGGFIFWRK--KRKDAKNNEVSFSQIKP----------KKSKLIRV 1023

Db        731 GLFLIGMLVALVAFFICRQKASHSRERPSARLSIRRDRPLSVHLNLGQKGNRKTSCPIKI 790

Qy       1024 ENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLPYDISRVK 1083

Db        791 NQFEGHFMKLQADSNYLLSKEYEDLKDVGRSQSCDIALLPENRGKNRYNNILPYDASRVK 850

Qy       1084 L-SVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAIIMLTKCV 1142

Db        851 LCNVDDDPCSDYINASYIPGNNFRREYIATQGPLPGTKDDFWKMAWEQNVHNIVMVTQCV 910

Qy       1143 EQGRTKCEEYWPSKQ-AQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHPL-RQFHF 1200

Db        911 EKGRVKCDHYWPADQDPLYYGDLILQMVSESVLPEWTIREFKICSEEQLDAHRLIRHFHY 970

Qy       1201 TSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDRLIYQIE 1260

Db        971 TVWPDHGVPETTQSLIQFVRTVRDYINRSPGAGPTVVHCSAGVGRTGTFVALDRILQQLD 1030

Qy       1261 NENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSKVDLIYQNTTAMTIY 1320

Db       1031 SKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCVRDVLRAKKLRN----EQENPLFPIY 1086

Qy       1321 ENLAP 1325

Db       1087 ENVNP 1091
```

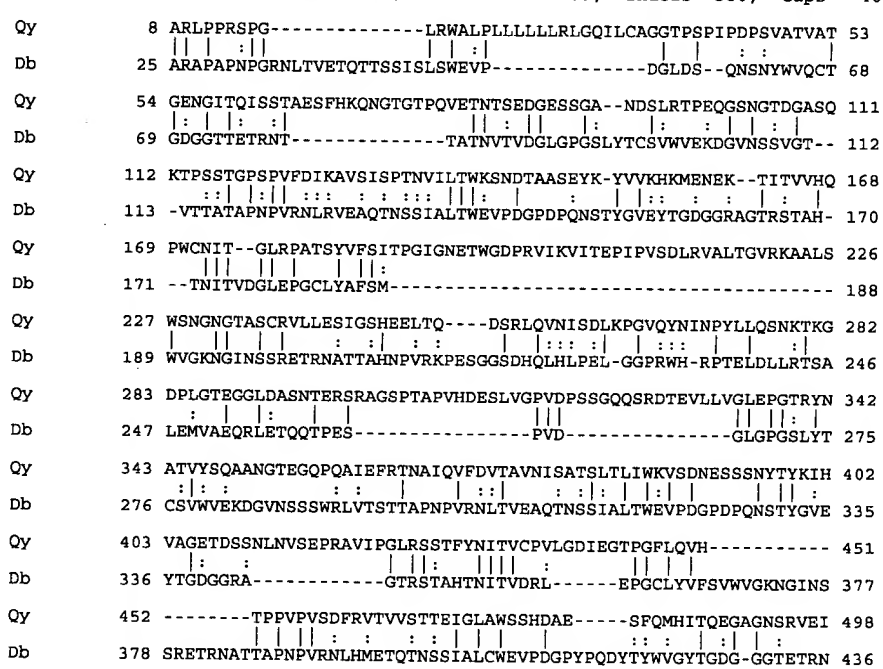
RESULT 14
Q9HD43
ID   Q9HD43   PRELIMINARY;   PRT; 1115 AA.
AC   Q9HD43;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transmembrane-type protein tyrosine phosphatase H.
GN   Name=PTPRH;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21328879; PubMed=11435690;
RA   Marneros A.G., Mehenni H., Reichenberger E., Antonarakis S.E.,
RA   Krieg T., Olsen B.R.;
RT   "Gene for the human transmembrane-type protein tyrosine phosphatase H
RT   (PTPRH): genomic structure, fine-mapping and its exclusion as a
RT   candidate for Peutz-Jeghers syndrome.";
RL   Cytogenet. Cell Genet. 92:213-216(2001).
DR   EMBL; AF275150; AAF91411.1; -.
DR   EMBL; AF275131; AAF91411.1; JOINED.
DR   EMBL; AF275132; AAF91411.1; JOINED.
DR   EMBL; AF275133; AAF91411.1; JOINED.
DR   EMBL; AF275134; AAF91411.1; JOINED.
DR   EMBL; AF275135; AAF91411.1; JOINED.
DR   EMBL; AF275136; AAF91411.1; JOINED.
DR   EMBL; AF275137; AAF91411.1; JOINED.
DR   EMBL; AF275138; AAF91411.1; JOINED.
DR   EMBL; AF275139; AAF91411.1; JOINED.
DR   EMBL; AF275140; AAF91411.1; JOINED.
DR   EMBL; AF275141; AAF91411.1; JOINED.
DR   EMBL; AF275142; AAF91411.1; JOINED.
DR   EMBL; AF275143; AAF91411.1; JOINED.
DR   EMBL; AF275144; AAF91411.1; JOINED.
DR   EMBL; AF275145; AAF91411.1; JOINED.
DR   EMBL; AF275146; AAF91411.1; JOINED.
DR   EMBL; AF275147; AAF91411.1; JOINED.
DR   EMBL; AF275148; AAF91411.1; JOINED.
DR   EMBL; AF275149; AAF91411.1; JOINED.
DR   HSSP; P18052; 1YFO.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 7.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00060; FN3; 7.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 7.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase; Transmembrane.
SQ   SEQUENCE   1115 AA; 122351 MW; A7A1AB24CFCD3846 CRC64;

  Query Match   16.7%; Score 1176.5; DB 2; Length 1115;
  Best Local Similarity   26.3%; Pred. No. 6.3e-58;
  Matches 366; Conservative 188; Mismatches 456; Indels 383; Gaps   43;

```
Qy          8 ARLPPRSPG--------------LRWALPLLLLLLRLGQILCAGGTPSPIPDPSVATVAT 53

Db         25 ARAPAPNPGRNLTVETQTTSSISLSWEVP--------------DGLDS--QNSNYWVQCT 68

Qy         54 GENGITQISSTAESFHKQNGTGTPQVETNTSEDGESSGA--NDSLRTPEQGSNGTDGASQ 111

Db         69 GDGGTTETRNT--------------TATNVTVDGLGPGSLYTCSVWVEKDGVNSSVGT-- 112

Qy        112 KTPSSTGPSPVFDIKAVSISPTNVILTWKSNDTAASEYK-YVVKHKMENEK--TITVVHQ 168

Db        113 -VTTATAPNPVRNLRVEAQTNSSIALTWEVPDGPDPQNSTYGVEYTGDGGRAGTRSTAH- 170

Qy        169 PWCNIT--GLRPATSYVFSITPGIGNETWGDPRVIKVITEPIPVSDLRVALTGVRKAALS 226

Db        171 --TNITVDGLEPGCLYAFSMWVG-KNGINSSRETRNATTAHNPVRNLRVEAQTTSSISLS 227

Qy        227 WSNGNGTASCRVLLESIGSHEELTQDSRLQVNISDLKPGVQYNINPYLLQSNKTKGDPLG 286

Db        228 WEVPDGTD--------------PQNS-----------------TYCVQCT-------G 247

Qy        287 TEGGLDASNTERSRAGSPTAPVHDESLVGPVDPSSGQQSRDTEVLLVGLEPGTRYNATVY 346

Db        248 DGGRTETRNT----------------------------TDTRVTVDGLGPGSLYTCSVW 278

Qy        347 SQAANGTEGQPQAIEFRTNAIQVFDVTAVNISA---TSLTLIWKVSDNESSSNYTYKIHV 403

Db        279 VE----KDGVNSSVEIVTSATAPNPVRNLTVEAQTNSSIALTWEVPDGPDPQNSTYGVEY 334

Qy        404 AGETDSSNLNVSEPRAVIPGLRSSTFYNITVCPVLGDIEGTPGFLQVH------------ 451

Db        335 TGDGGRA------------GTRSTAHTNITVDRL------EPGCLYVFSVWVGKNGINSS 376

Qy        452 -------TPPVPVSDFRVTVVSTTEIGLAWSSHDAE-----SFQMHITQEGAGNSRVEIT 499

Db        377 RETRNATTAPNPVRNLHMETQTNSSIALCWEVPDGPYPQDYTYWVEYTGDG-GGTETRNT 435
```



```
RESULT 15
Q15426
ID   Q15426      PRELIMINARY;      PRT;  1118 AA.
AC   Q15426;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Protein tyrosine phosphatase precursor.
GN   Name=SAP-1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94124561; PubMed=8294459;
RA   Matozaki T., Suzuki T., Uchida T., Inazawa J., Ariyama T., Matsuda K.,
RA   Horita K., Noguchi H., Mizuno H., Sakamoto C., Kasuga M.;
RT   "Molecular cloning of a human transmembrane-type protein tyrosine
RT   phosphatase and its expression in gastrointestinal cancers.";
RL   J. Biol. Chem. 269:2075-2081(1994).
DR   EMBL; D15049; BAA03645.1; -.
DR   PIR; A49724; A49724.
DR   HSSP; P18052; 1YFO.
DR   Genew; HGNC:9672; PTPRH.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005001; F:transmembrane receptor protein tyrosine pho. . .; TAS.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00041; fn3; 6.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   SMART; SM00060; FN3; 6.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS50853; FN3; 6.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase; Signal.
FT   SIGNAL        1     25       Potential.
FT   CHAIN        26   1118       protein tyrosine phosphatase.
SQ   SEQUENCE   1118 AA;  123038 MW;  F73E3967ECE59F00 CRC64;

  Query Match           16.1%;  Score 1130;  DB 2;  Length 1118;
  Best Local Similarity  25.9%;  Pred. No. 2.7e-55;
  Matches  354;  Conservative  181;  Mismatches  453;  Indels  380;  Gaps   40;
```



```
Qy        499 TTNQSIIIGGLFPGTKYCFEIVPKGPNGTEGASRTVCNRTVPSAVFDIHVVYVTTTEMWL 558
Db        437 TTNTSVTAERLEPGTLYTFSVWAE-KNGARGSRQNVSISTVPNAVTSLSKQDWTNSTIAL 495

Qy        559 DWKSPD--GASEYVYHL--VIESKHGSNHTSTYDKAITLQGLIPGTLYNITISPEVDHVW 614
Db        496 RWTAPQGPGQSSYSYWVSWVREGMTDPRTQSTSGTDITLKELEAGSLYHLTVWAERNEVR 555

Qy        615 GDPNSTAQYTRPSNVSNIDVSTNTTAATLSWQNF--DDASPTYSYCLLIEKAGN------ 666
Db        556 GYNSTLTAATAPNEVTDLQNETQTKNSVMLWWKAPGDPHSQLYVYWVQWASKGHPRRGQD 615

Qy        667 -SSNATQVVTDIGITDATVTELIPGSSYTVEIFAQVGDGIKSLEPGRKSFC--TDPASMA 723
Db        616 PQANWVNQTSRTNETWYKVEALEPGTLYNFTVWAERNDVASSTQ----SLCASTYPDTVT 671

Qy        724 SFDCEVVPKEPALVLKWTCPPGANAGFELEVSSGAWNNATHLESCSSENGTEYRTEVTYL 783
Db        672 ITSCVSTSAGYGVNLIWSCPQGGYEAFELEV--------------GGQRGSQDRS----- 712

Qy        784 NFSTSYNISITTVSCGKMAAPTRNTCTTGITDPPPPDGSPNITSVSHNSVKVKFSGFEAS 843
Db        713 -------------SCGEAVS------VLGL------------------------------ 723

Qy        844 HGPIKAYAVILTTGEAGHPSADVLKYTYDDFKKGASDTYVTYLIRTEEKGRSQSLSEVLK 903
Db        724 -GPARSYPATITT-------------IWDGMK------------------------VVS 744

Qy        904 YEIDVGNESTTLGYYNGKLEPLGSYRACVAGFTNITFHPQNKGLIDGAESYVSFSRYSDA 963
Db        745 HSVVCHTESA---------------------------------------------------- 754

Qy        964 VSLPQDPGVICGAVFGCIFGALVIVTVGGFIFWRKKRKDAKNNEVSFSQIKPKKSKL--- 1020
Db        755 -------GVIAGAFVGIL---LFLILVGLLIFFLKRRNKKK-------QQKPELRDLVFS 797

Qy       1021 ----IRVENPEAYFKKQQADSNCGFAEEYEDLKLVGISQPKYAAELAENRGKNRYNNVLP 1076
Db        798 SPGDIPAEDFADHVRKNERDSNCGFADEYQQLSLVGHSQSQMVASASENNAKNRYRNVLP 857

Qy       1077 YDISRVKLS-VQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTLKDFWRMVWEKNVYAI 1135
Db        858 YDWSRVPLKPIHEEPGSDYINASFMPGLWSPQEFIATQGPLPQTVGDFWRLVWEQQSHTL 917

Qy       1136 IMLTKCVEQGRTKCEEYWP-SKQAQDYGDITVAMTSEIVLPEWTIRDFTVKNIQTSESHP 1194
Db        918 VMLTNCMEAGRVKCEHYWPLDSQPCTHGHLRVTLVGEEVMENWTVRELLLLQVEEQKTLS 977

Qy       1195 LRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGVGRTGTFIAIDR 1254
Db        978 VRQFHYQAWPDHGVPSSPDTLLAFWRMLRQWLDQTMEGGPPIVHCSAGVGRTGTLIALDV 1037

Qy       1255 LIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQ 1302
Db       1038 LLRQLQSEGLLGPFSFVRKMRESRPLMVQTEAQYVFLHQCICGSSNSQ 1085
```

Search completed: May 10, 2005, 19:04:41
Job time : 218 secs